PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/705, C12N 15/81, 1/19, C12Q 1/02 | A1

(11) International Publication Number: WO 96/13520

(43) International Publication Date: 9 May 1996 (09.05.96)

(21) International Application Number: PCT/US95/14364

(22) International Filing Date: 25 October 1995 (25.10.95)

(30) Priority Data:
08/332,312 31 October 1994 (31.10.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/332,312 (CIP)
Filed on 31 October 1994 (31.10.94)

(71) Applicant *(for all designated States except US)*: AMERICAN CYANAMID COMPANY [US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PAUSCH, Mark, Henry [US/US]; 312 Andover Place, Robbinsville, NJ 08692 (US). PRICE, Laura, Alicia [US/US]; 181 Canterbury Avenue, Langhorne, PA 19047 (US).

(74) Agents: MATHEWS, Gale, F.; American Home Products Corporation, Five Giralda Farms, Madison, NJ 07940 (US) et al.

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SD, SG, SI, SK, TJ, TM, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: GENES ENCODING A FAMILY OF POTASSIUM CHANNELS

(57) Abstract

This invention relates generally to the potassium channel gene family. More particularly, the present invention relates to the cloning and characterization of potassium channel genes from *Drosophila melanogaster* and *Caenorhabditis elegans*. Other aspects of the present invention include methods of assaying substances to determine effects on cell growth. Also presented are methods of controlling nematode and insect pests by inhibiting potassium channels substantially homologous to those encoded by nucleotide sequences as described herein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

32,421-01

GENES ENCODING A FAMILY OF POTASSIUM CHANNELS

Field of Invention

This invention relates generally to the potassium channel gene family. More particularly, the present invention relates to the cloning and characterization of potassium channel genes from *Drosophila melanogaster* and *Caenorhabditis elegans*.

Background of the Invention

Synthetic organic insecticides are primarily nerve poisons acting on the cholinergic system (organophosphorus compounds and methylcarbamates), the voltage-gated sodium channel (pyrethroids and DDT), and the GABA-gated chloride channel (cyclodienes and other polychlorocycloalkanes). Potassium channels comprise a large and diverse group of integral membrane proteins that determine the level of excitability and repolarization properties of neurons and muscle fibers [B. Hille, *Ionic Channels of Excitable Membranes*, Sinauer, Sunderland, MA (1984)]. The multiple essential functions encoded by the potassium channels make them excellent targets for new pesticides and animal and human therapeutics. Potassium channel diversity in the fruitfly *Drosophila melanogaster* results from an extended gene family coding for homologous proteins. Six genes encoding

potassium channels have been cloned from *Drosophila melanogaster* which account for a large part of the diversity of potassium currents observed in insect nervous tissue [A. Wei, M. Covarrubias, A. Butler, K. Baker, M. Pak, L. Salkoff, *Science* 248, 599-603 (1990), N.S. Atkinson, G.A. Robertson, B. Ganetzky, *Science* 253,551-555, (1991), J. Warmke, R. Drysdale, B. Ganetzky, *Science* 252, 1560-1564 (1991), A. Bruggemann, L.A. Pardo, W. Stuhmer, O. Pongs, *Nature* 365, 445-448 (1993)]. *Shaker* and *Shal* encode voltage-gated potassium channels with rapid current activation and inactivating properties. *Shab* and *Shaw* encode delayed rectifier channels, with slow inactivating (*Shab*) and non-inactivating (*Shaw*) properties. *Slo* encodes a calcium-activated potassium channel and *eag* encodes a voltage-gated channel permeable to both potassium and calcium which is modulated by cyclic AMP.

Modulation of cardiac action potential by compounds that effect the behavior of potassium channels may be a useful treatment for serious heart conditions. In this regard, each of the potassium channels cloned from insects have corresponding versions in mammalian species, including, specifically, a delayed rectifier potassium channel homolog, RAK, cloned from rat cardiac tissue [M. Paulmichl, P. Nasmith, R. Hellmiss, K. Reed, W.A. Boyle, J.M. Nerbonne, E.G. Peralta, D.E. Clapham, *Proc. Natl. Acad. Sci USA* 88, 7892-7895 (1991)]. Thus, the RAK channel represents an important target of new drugs for the control of heart failure. The d layed rectifier potassium current in heart cells regulates the duration of the plateau of the cardiac action potential by countering the depolarizing,

inward calcium current. Delayed rectifier potassium currents characteristically are activated upon depolarization from rest, display a sigmoidal or delayed onset, and have a nonlinear, or rectifying, current-voltage relation. Several types of delayed potassium conductances have been identified in cardiac cells based on measured single-channel conductances. Heart rate and contractility are regulated by second messenger modification of delayed rectifier potassium conductances, and species differences in the shape of the plateau may be influenced by the type and level of channel expression.

On the basis of predicted membrane spanning topology, potassium channels may be subdivided into two distinct classes: voltage-gated, calcium-activated, and cyclic nucleotide-gated potassium channels that are composed of six membrane spanning domains (S1-S6) and a single pore forming domain (H5), and inward rectifying potassium channels that pass through the membrane twice and also contain a single pore forming region [Y. Kubo, E. Reuveny, P.A. Slesinger, Y.N. Jan, L.Y. Jan *Nature* 364, 802-806 (1993); Y. Kubo, T.J. Baldwin, Y.N. Jan, L.Y. Jan *Nature* 362, 127-133 (1993)]. Here, we report the cloning and functional expression in yeast of a novel *Drosophila melanogaster* potassium channel. Further, we identify a *Caenorhabditis elegans* homolog that constitutes the second member of a new family of potassium channels exhibiting a topological configuration unique among the known classes of potassium channels.

The yeast *Saccharomyces cerevisiae* is utilized as a model eukaryotic organism for the purpose of studying potassium transport mechanisms.

Due to the ease with which one can manipulate the genetic constitution of the yeast *Saccharomyces cerevisiae*, researchers have developed a detailed understanding of many complex biological pathways, including potassium transport. In yeast, high affinity potassium uptake is performed by the product of the *TRK1* gene [R.F. Gaber, C.A. Styles, G.R. Fink *Mol. Cell. Biol.* 8, 2848-2859 (1988)]. Mutant yeast strains lacking *trk1* function are incapable of growing in medium lacking high concentrations of potassium. Since potassium transport mechanisms are present in organisms as divergent as yeast and man, one could predict that expression of heterologous potassium channels in mutant cells might replace *trk1* function, and support growth on medium containing low potassium concentration. In this regard, plant potassium channels were shown to function in yeast and represent important targets for new herbicides [J.A Anderson, S.S. Huprikar, L.V. Kochian, W.J. Lucas, R.F. Gaber, *Proc. Natl. Acad. Sci USA* 89, 3736-3740 (1992); H. Sentenac, N. Bonnaud, M. Minet, F. Lacroute, J.-M. Salmon, F. Gaynard, C. Grignon, *Science* 256, 663-665 (1992); D.P. Schachtman and J.I. Schroeder, *Nature* 370, 655-658]. Thus, we have employed this yeast expression system for cloning and expression of potassium channels from heterologous species, making it useful for discovery of new pesticides, and animal and human therapeutics. Discovery of such compounds will necessarily require screening assays of high specificity and throughput. For example, new pesticides directed at potassium channels require high selectivity for insect channels and low activity against non-insect species. Screening assays utilizing yeast strains genetically modified to

accommodate functional expression of heterologous potassium channels offer significant advantages in this area.

Summary of the Invention

A first aspect of the present invention is the discovery of a new subclass of potassium channel genes and proteins encoded thereby. Potassium channels belonging to this new subclass comprise four hydrophobic domains capable of forming transmembrane helices, wherein a first pore-forming domain is interposed between the first and second transmembrane helices and a second pore-forming domain is interposed between the third and fourth transmembrane helices, and wherein each pore-forming domain contains a potassium selective peptide motif. In preferred embodiments, the peptide motif is selected from the group consisting of a Y/F-G dipeptide motif.

In certain preferred embodiments, the isolation and characterization of invertebrate (i.e. insect and nematode) potassium channel genes belonging to this new subclass is presented. In more preferred embodiments, the present invention provides for the isolation of complementary DNA fragments from *Drosophila melanogaster* and *Caenorhabditis elegans* which encode conserved amino acid sequence elements unique to this potassium channel gene family. A yeast expression technology is employed to clone cDNAs from *Drosophila melanogaster* and *C. elegans* and a hybridization approach is utilized to isolate additional cDNAs from *Caenorhabditis elegans*.

A second aspect of the present invention is a method of assaying substances to determine effects

on cell growth. Yeast cells of the kind described above are cultured in appropriate growth medium to cause expression of heterologous proteins, embedded in agar growth medium, and exposed to chemical compounds applied to the surface of the agar plates. Effects on the growth of embedded cells are found around compounds that have effects on the heterologous potassium channel.

A third aspect of the present invention is a method of controlling nematode and insect pests by inhibiting potassium channels substantially homologous to those encoded by nucleotide sequences as presented herein.

Brief Description of the Drawings

FIGURE 1. Growth of CY162 cells bearing pDmORF1. CY162 cells transformed with plasmids isolated from survivors of a primary library screen for plasmids that support the growth of CY162 on medium contain low potassium concentration. Six individual transformants of each plasmid-bearing strain are cultured in patches on the indicated medium. CY162 cells bearing pDmORF1 are found in the upper left-hand corner of each plate while pKAT1 containing cells are found in the lower right hand corner.

FIGURE 2A and 2B. DNA sequence and deduced amino acid sequence of Dm ORF1 [SEQ ID NOS:1 and 2]. The nucleotide sequence of the 2.4 kb cDNA revealed a single long open reading frame pr ximal to the GAL1 promoter. Segments corresponding to putative transmembrane (M1-M4) and pore-f rming H5 domains in the predicted polypeptide are underlined. The single

amino-terminal asparagine linked glycosylation site is indicated by a G.

FIGURE 3A and 3B. DNA sequence and deduced amino acid sequence of the F22b7.7 segment of the *Caenorhabditis elegans* genome [SEQ ID NO:3]. Segments corresponding to putative transmembrane (M1-M4) and pore-forming H5 domains in the predicted polypeptide are underlined.

FIGURE 4. Alignment of DmORF1 and F22b7.7 sequences. Protein-coding regions of DmORF1 [SEQ ID NO: 37] and F22b7.7 [SEQ ID NO: 38] (designated as CeORF-1 in this FIGURE) are compared using the protein sequence alignment algorithm in Genework DNA sequence analysis software. Identical amino acids are boxed.

FIGURE 5A. Comparison of the pore-forming domains of DmORF1 and F22b7.7. Amino acid sequences from the six cloned *Drosophila melanogaster* potassium channels and three inward rectifier channels [SEQ ID NOS:7 through 21] are compared to DmORF1 and F22b7.7 within the pore-forming H5 regions. Amino acid identities are indicated by a vertical line and conserved substitutions indicated by a dot. Amino acid substitutions deemed acceptable are indicated.

FIGURE 5B. Hydropathy plot analysis of the DmORF1 and F22b7.7 polypeptide sequence. The Kyte-Doolittle hydropathy algorithm in the Geneworks DNA analysis software is used to predict the topology of DmORF1 and F22b7.7. The position of predicted membrane spanning domains (M1-M4) and pore-forming domains are indicated.

FIGURE 6. Predicted membrane spanning topology of DmORF1.

FIGURE 7. Heterologous potassium channel-dependent growth of plasmid bearing CY162 (*trk1Δ*) strains. CY162 bearing pYES2, pKAT1, pDmORF1, and pRATRAK are cultured at 30°C for four days on arginine phosphate agar medium containing 0 mM, 0.2 mM, or 100 mM added KCl.

FIGURE 8. Inhibition of growth of yeast cells containing heterologous potassium channels. CY162 cells (10^5) bearing the indicated plasmids are plated in arginine phosphate agar medium containing 0.2 mM potassium chloride. Sterile filter disks were placed on the surface of the agar and saturated with 20 ml of a 1 M solution of potassium channel blocking compound. Clockwise from upper left-hand corner is $BaCl_2$, CsCl, TEA, and RbCl. KCl is applied to the center disk.

FIGURE 9A and 9B. DNA sequence and deduced amino acid sequence of CORK [SEQ ID NO: 36]. The nucleotide sequence of the 1.4 kb cDNA revealed a single long open reading frame proximal to the *GAL1* promoter. Segments corresponding to pore-forming H5 domains in the predicted polypeptide are underlined. Asparagine-linked glycosylation sites are indicated by a G.

Detailed Description of the Invention

Nucleotide bases are abbreviated herein as follows:
Ade;A-Adenine G-Guanine Ura; U-Uracil
C-Cytosine T-Thymine
Amino acid residues are abbreviated herein to either three

letters or a single letter as follows:
Ala;A-Alanine Leu;L-Leucine
Arg;R-Arginine Lys;K-Lysine
Asn;N-Asparagine Met;M-Methionine
Asp;D-Aspartic acid Phe;F-Phenylalanine
Cys;C-Cysteine Pro;P-Proline
Gln;Q-Glutamine Ser;S-Serine
Glu;E-Glutamic acid Thr;T-Threonine
Gly;G-Glycine Trp;W-Tryptophan
His;H-Histidine Tyr;Y-Tyrosine
Ile;I-Isoleucine Val;V-Valine

The term "mammalian" as used herein refers to any mammalian species (e.g., human, mouse, rat, and monkey).

The term "heterologous" as used herein refers to DNA sequences, proteins, and other materials originating from organisms other than the organism used in the expression of the potassium channels or portions thereof, or described herein (e.g., mammalian, avian, amphibian, insect, plant), or combinations thereof not naturally found in yeast.

The terms "upstream" and "downstream" are used herein to refer to the direction of transcription and translation, with a sequence being transcribed or translated prior to another sequence being referred to as "upstream" of the latter.

The potassium channels of the present invention possess properties in common with known potassium channels including, voltage-gated channels, calcium activated channels, cyclic nucleotide gated channels, inward rectifier channels, and the like, and especially with regard to electrophysiological properties. Certain preferred channels exhibit inward and outward currents that are affected by potassium concentration, particularly characteristic of voltage-gated channels. The term "channel" and the

nucleotide sequences encoding same, is intended to encompass subtypes of the aforementioned classes of channels, and mutants, derivatives and homologs thereof.

The nucleotide sequences encoding the potassium channels or parts thereof may be expressed recombinantly, and utilized for a variety of reasons, the most notable of which is for screening of substances that modulate the activity of the potassium ion channels. Such substances, especially inhibitors of the activity of the potassium channels of the present invention, may be utilized as insecticides, antihelmenthics, drugs suitable for the control of heart failure, and the like.

Heterologous DNA sequences are typically expressed in a host by means of an expression vector. An expression vector is a replicable DNA construct in which a DNA sequence encoding the heterologous DNA sequence is operably linked to suitable control sequences capable of affecting the expression of a protein or protein subunit coded for by the heterologous DNA sequence in the intended host. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and (optionally) sequences which control the termination of transcription and translation. Vectors useful for practicing the present invention include plasmids, viruses (including bacteriophage), and integratable DNA fragments (i.e., fragments integratable into the host genome by genetic recombination). The vector may replicate and function independently of the host genome, as in the case of a plasmid, or may integrate into the genome itself, as in the case of an integratable DNA fragment. Suitable vectors will contain replicon and control sequences which are derived from species compatible with the intended expression host. For example, a promoter operable in a host cell is

one which binds the RNA polymerase of that cell, and a ribosomal binding site operable in a host cell is one which binds the endogenous ribosomes of that cell.

DNA regions are operably associated when they are functionally related to each other. For example, a promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation. Generally, operably linked means contiguous and, in the case of leader sequences, contiguous and in reading phase.

Transformed host cells of the present invention are cells which have been transformed or transfected with the vectors constructed using recombinant DNA techniques and express the protein or protein subunit coded for by the heterologous DNA sequences. In preferred embodiments, the transformed host cells are yeast. A variety of yeast cultures, and suitable expression vectors for transforming yeast cells, are known. See e.g., U.S. Patent No. 4,745,057; U.S. Patent No. 4,797,359; U.S. Patent No. 4,615,974; U.S. Patent No. 4,880,734; U.S. Patent No. 4,711,844; and U.S. Patent No. 4,865,989. *Saccharomyces cerevisiae* is the most commonly used among the yeasts, although a number of other yeast species are commonly available. See. e.g., U.S. Patent No. 4,806,472 (*Kluveromyces lactis* and expression vectors therefore); 4,855,231 (*Pichia pastoris* and expression vectors therefore). A heterologous potassium channel may permit a yeast strain unable to grow in medium containing low potassium concentration to survive [CY162, for example, see J.A Anderson, S.S. Huprikar, L.V. Kochian, W.J. Lucas, R.F. Gaber, *Proc. Natl. Acad. Sci USA* 89, 3736-3740 (1992)]. Yeast vectors may contain an origin of replication from the endogenous 2 micron (2μ) yeast plasmid or an autonomously

replicating sequence (ARS) which confer on the plasmid the ability to replicate at high copy number in the yeast cell, centromeric (CEN) sequences which limit the ability of the plasmid to replicate at only low copy number in the yeast cell, a promoter, DNA encoding the heterologous DNA sequences, sequences for poly-adenylation and transcription termination, and a selectable marker gene. An exemplary plasmid is YRp7, (Stinchcomb et al., (1979) *Nature* 282, 39; Kingsman et al., (1979) *Gene* 7, 141; Tschemper et al., (1980) *Gene* 10, 157]. This plasmid contains the TRP1 gene, which provides a selectable marker for a mutant strain of yeast lacking the ability to grow in the absence tryptophan, for example ATCC No. 44076. The presence of the trp1 lesion in the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for metallothionein (YEp52), 3-phosphoglycerate kinase [pPGKH, Hitzeman et al., (1980) *J. Biol. Chem.* 255, 2073] or other glycolytic enzymes [pYSK153, Hess et al., (1968) *J. Adv. Enzyme Reg.* 7, 149]; and Holland et al., (1978) *Biochemistry* 17, 4900], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phospho-fructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, trioseposphate isomerase, phosphoglucose isomerase, and glucokinase. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., EPO Publn. No. 73,657. Other promoters, which have the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydr genase 2 (pAD4M), isocytochrome C, acid phosphates, degradative enzymes associated with nitrogen metabolism, and the aforementioned metallothionein and glyceraldehyde-3-

phosphate dehydrogenase, as well as enzymes responsible for maltose and galactose (pYES2) utilization. Finally, in constructing suitable expression plasmids, the termination sequences associated with these genes may also be ligated into the expression vector 3' of the heterologous coding sequences to provide polyadenylation and termination of the mRNA.

In one embodiment of the present invention, a yeast expression system is described, wherein yeast cells bear heterologous potassium channels. In preferred embodiments, these channels are DmORF-1, CORK, or RAK. As noted above, transformed host cells of the present invention express the proteins or proteins subunit coded for by the heterologous DNA sequences. When expressed, the potassium channel is located in the host cell membrane (i.e., physically positioned therein in proper orientation for both the stereoselective binding of ligands and passage of potassium ions).

In certain preferred screening embodiments of the present invention, a transformed yeast cell is presented, containing a heterologous DNA sequence which codes for a rat cardiac delayed rectifier potassium channel, RAK, cloned into a suitable expression vector. RAK is capable of complementing the potassium-dependent phenotype of *Saccharomyces cerevisiae* strain CY162 on medium containing low potassium concentration.

The potassium channel subclass of the present invention is characterized in that the potassium channels have four hydrophobic domains capable of forming transmembrane helices. These channels are further characterized in that they comprise two pore-forming d mains, one of which is interposed between said first helix and said second helix, and the other of which is interposed between said third helix and said fourth helix. The pore-forming domains

 SUBSTITUTE SHEET (RULE 26)

further contain a potassium selective motif which serves to confer upon the channel the ability to pass potassium ions to the exclusion of other ions, such as sodium, calcium, and the like. In certain preferred embodiments, this motif contains the peptide Y/G, and particularly in either a dipeptide or tripeptide motif, and frequently with Y/F-G bonding. In most preferred embodiments, the motif is selected from the group consisting of G-V-G, G-L-G, G-Y-G, G-F-G, and G-I-G.

In certain embodiments of the present invention, the potassium channel is positioned within a cell membrane in such a manner as to allow it to function as a modulator of the flow of potassium ions into and out of the cell. To best regulate this activity, at least one pore-forming domain may be positioned proximal to a exterior portion of the cell membrane.

In other embodiments, the potassium channels of the present invention further comprise an amino-terminal glycosylation site, and especially wherein that site is asparagine-linked.

Potassium channels belonging to the subclass as presented herein may be derived from a wide variety of animal species, both vertebrate and invertebrate. Using the yeast expression technology and other teachings as set forth herein, the present inventors have isolated a single 2463 base pair cDNA fragment from an invertebrate source, designated Dm ORF1 [SEQ ID NO: 1], by complementation of the potassium-dependent phenotype of *Saccharomyces cerevisiae* strain CY162 (*trk1Δ*) on medium containing low potassium concentration [J.A Anderson, S.S. Huprikar, L.V. Kochian, W.J. Lucas, R.F. Gaber, *Proc. Natl. Acad. Sci USA* 89, 3736-3740 (1992)]. Dm ORF1 contains a single long open reading frame encoding a protein of 618 amino acids [SEQ ID NO:2] that exhibits substantial amino acid identity to the pore-

forming regions of other potassium channels. The DmORF1 contains structural features that distinguish it from other classes of potassium channels, including four hydrophobic domains capable of forming transmembrane helices (M1-M4) and two putative pore forming H5 domains found between transmembrane helices M1 and M2, and M3 and M4. Each pore forming H5 domain contains the Y/F-G dipeptide motif required for potassium selectivity [L. Heginbotham, T. Abramson, R. MacKinnon, *Science* 258, 1152-1155, (1992)]. This work was expanded to clone a construct derived from *C. elegans* having a single open reading frame sufficient to encode a protein of 434 amino acids, designated pCORK.

A search of the GENBANK database for DNA and protein sequences similar to DmORF1 revealed several cloned potassium channel sequences including a putative protein coding DNA sequence, F22b7.7, reported in the *Caenorhabditis elegans* genome sequencing project [R. Wilson, R. Ainscough, K Anderson, *et al. Nature* 368, 32-38 (1994)]. The DNA sequence contained a single long open reading frame sufficient to encode a protein of 336 amino acids (predicted MW 38.5 kDa) with substantial homology to known potassium channel sequences.

Using the hybridization approach, a cDNA sequence designated CeORF1 [SEQ ID NO: 38] was isolated by probing a *Caenorhabditis elegans* cDNA library with oligonucleotides designed using F22b7.7 DNA sequences [T.N. Davis and J. Thorner *Meth. Enzymol.* 139, 246-262 (1987)]. CeORF1 contains a single long open reading frame encoding a protein that exhibits substantial amino acid identity to pore-forming regions of other potassium channels.

CeORF1 and pCORK each contain structural features similar to DmORF1, including two putative pore forming H5 domains. Each pore forming H5 domain contains the Y/F-G dipeptide motif required for potassium selectivity [L.

Heginbotham, T. Abramson, R. MacKinnon, *Science* 258, 1152-1155, (1992)]. These features form the basis of the designation of a new sub-family of potassium channels comprising DmORF1, CORK, and CeORF1.

Other aspects of the present invention relate to methods of modulating potassium channel activity, by affecting the ability of such channel to allow the flow of ions into, through, or out of a cellular membrane, and particularly when these ions are potassium ions. Certain substances whether biological or chemical in nature, may be applied to cell membranes having as an integral part of their structure, one or more potassium channels comprising the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 36, or RAK, in an amount and for a time sufficient to affect the ability of the potassium channel to so regulate the flow of ions. Substances that are potassium channel blockers will inhibit the ability of the channel to regulate the flow of such ions. Substances that enhance such ability may be considered potassium channel "activators." Substances that modulate the activity of RAK may do so by modulation of cardiac action potential, upward or downward.

Application of such substances may take the form of *in vitro, ex vivo,* or *in vivo* application, each in a formulation suitable to deliver the substance to the cell membrane and to sustain such delivery for a time sufficient to allow the substance to interact with the membrane. Appropriate formulations, concentrations of substances, application time, and other relevant parameters may be established by utilizing, *inter alia*, known assays for measuring ion channel current flow. Another suitable endpoint one skilled in the art may utilize in optimizing these parameters, especially in the case of potassium channel blockers, is "cell death". Such assays may be performed *in vitro* and extrapolated to *in vivo* conditions,

or in some cases may be easily established directly *in vivo*, as for example, by applying the substance directly to a test sample comprising the target insect pest (whole organism) and noting the appropriate parameters at which an acceptable per cent of insect death is attained.

In certain preferred embodiments, methods of selectively inhibiting insect pests are presented by applying to such insect pests a substance capable of selectively inhibiting the activity of a potassium channel contained in the cells of such insect, and comprising the amino acid sequence of SEQ ID NO:2, or a potassium channel substantially homologous thereto. In the most preferred embodiments, the inhibitor will inhibit the activity of the aforementioned potassium channel without inhibition of other, non-homologous potassium channels that may be present in species other than the targeted insect pest. It is envisioned that such other species may also be present at the site of application of the inhibitor, such as in a garden, crop, or other site wherein it is desired to control insect pests. In other preferred embodiments, methods of selectively inhibiting nematode pests are presented much in the same manner as discussed for control of insect pests, by applying to such pests a substance capable of selectively inhibiting the activity of a potassium channel contained in the cells of such pest, and comprising the the amino acid sequence of SEQ ID NO:4 or SEQ ID NO: 36, or potassium channels substantially homologous thereto.

The following Examples are provided to further illustrate various aspects of the present invention. They are not to be construed as limiting the invention.

Example 1

Recombinant expression library screening.

Saccharomyces cerevisiae strain CY162 is described in Anderson, J.A. *et al.* (1992) *Proc. Natl. Adad. Sci. USA* 89, 3736-3740]. Growth of bacterial strains and plasmid manipulations are performed by standard methods (Maniatis T., Molecular Cloning. Cold Spring Harbor Laboratory Press, 1982). Media conditions for growth of yeast, isolation of plasmid DNA from yeast, and DNA-mediated transformation of yeast strains are as described (Rose M. D., Methods in yeast genetics, Cold Spring Harbor Laboratory Press, 1990). A multifunctional expression library constructed in pYES2 and containing cDNA made from 3rd instar male *Drosophila melanogaster* mRNA is used as described [S.J. Elledge, J.T. Mulligan, S.W. Ramer, M. Spottswood, R.W. Davis *Proc. Natl. Acad. Sci USA* 88, 1731-1735 (1991)]. A multifunctional expression library constructed in pYES2 and containing cDNA made from mRNA obtained from all life stages of *Caenorhabditis elegans* is custom-made by Invitrogen Corporation.

Isolation of expression plasmids encoding heterologous potassium channels. CY162 cells are transformed with plasmid DNA from each library to give 3 x 10^6 transformants from each library on SCD-ura (synthetic complete dextrose (2 %) medium containing all necessary nutritional supplements except uracil) containing 0.1 M KCl agar medium. Transformants are replica-plated to SCG-ura (synthetic complete galactose (2 %) medium containing all necessary nutritional supplements except uracil) agar medium. Colonies that grow on this selective agar medium are transferred to SCG-ura agar medium to obtain single colonies clones and while reassaying suppression of the potassium-dependent phenotype. Plasmid DNA is isolated from surviving colonies and used to transform CY162. Six individual transformant strains containing one plasmid, pDmORF1,that confers the potassium independent phenotype is cultured on

SCD-ura and SCG-ura medium along with CY162 strains bearing pKAT1, which encodes a plant inward rectifier potassium channel that supports the growth of CY162 on selective medium (FIGURE 1). The plasmid bearing strains exhibit potassium-independent growth on both dextrose and galactose containing medium. Growth on dextrose is likely due to basal level of transcription leading to sufficient potassium channel expression to support growth.

Example 2

DNA sequence analysis of DmORF1. Plasmids that confer suppression of the potassium-dependent phenotype are subjected to automated DNA sequence analysis performed by high temperature cycle sequencing (Applied Biosystems). Geneworks DNA sequence analysis software (Intelligenetics) is used to align raw DNA sequence information and to identify open reading frames. The DNA sequence of the 2.4 kb insert in pDmORF1 is displayed in FIGURE 2A and 2B [SEQ ID NO:1]. The 5' untranslated sequences of the cDNA contain long poly A and poly T tracts not likely to be found in protein coding regions. The first ATG proximal to the 5' end is present in a consensus *Drosophila melanogaster* translational initiation site [D.R. Cavener *Nucleic Acids Res.*, 15, 1353-1361 (1987)], consistent with the designation of this site as the translational start site. A single long open reading frame sufficient to encode a protein of 618 amino acids (predicted MW 68 kDa) is encoded in pDmORF1. A consensus polyadenylation site, AATCAA, occurs at position 2093-2098 in 3' untranslated sequences. The DmORF1 contains structural features that distinguish it fr m other classes of potassium channels, including four hydroph bic domains capable of forming transmembrane helices (M1-M4) and two pore forming H5 domains found between transmembrane helices

M1 and M2, and M3 and M4. Each pore forming H5 domain contains the Y/F-G dipeptide motif required for potassium selectivity [L. Heginbotham, T. Abramson, R. MacKinnon, *Science* 258, 1152-1155, (1992)].

Example 3

Identification of *Caenorhabditis elegans* sequences homologous to DmORF1. A search of the GENBANK database protein sequences similar to DmORF1 reveals significant matches with several known potassium channel sequences. The closest match is to a putative protein coding DNA sequence, F22b7.7, reported in the *Caenorhabditis elegans* genome sequencing project [R. Wilson, R. Ainscough, K. Anderson, *et al.*, *Nature* 368, 32-38 (1994)]. The DNA sequence and predicted amino acid sequence assembled from putative exons recognized by a GENBANK exon identification algorithm is displayed in FIGURE 3A and 3B [SEQ ID NOS:3 and 4]. The DNA sequence contains a single long open reading frame sufficient to encode a protein of 336 amino acids (predicted MW 38.5 kDa) with substantial homology to known potassium channel sequences. The F22b7.7 sequence contains structural features that distinguish it from other classes of potassium channels, including three of four hydrophobic domains capable of forming transmembrane helices (M1-M4) identified in DmORF1 and two pore forming H5 domains found between transmembrane helices a predicted M1 and M2, and M3 and M4. Each pore forming H5 domain contains the Y/F-G dipeptide motif required for potassium selectivity [L. Heginbotham, T. Abramson, R. MacKinnon, *Science* 258, 1152-1155, (1992)]. The lack of an amino terminal transmembrane domain homologous to DmORF1 M1 in the F22b7.7 sequence may be due to failure of the search algorithm to identify exon(s)

encoding the amino terminus. Alternatively, an amino terminal coding sequence may be added by trans-splicing, which occurs frequently in *Caenorhabditis elegans*.

Example 4

Cloning and DNA sequence analysis of CeORF1. Oligonucleotides corresponding to DNA sequences encoding the two pore forming domains of F22b7.7 are synthesized using an Applied Biosystems DNA synthesizer. F22b7.7-H2-1: 5'TCCATTTTCTTTGCCGTAACCGTCGTCACTACCATCGGATACGGTAATCCA [SEQ ID NO:5]. F22b7.7-H2-2: 5'TCATTCTACTGGTCCTTCATTACAATGACTACTGTCGGGTTTGGCGACTTG [SEQ ID NO:6]. The oligos were labelled at their 5' ends with ^{32}P using a 5'-end labelling kit according to manufacturers instructions (New England Nuclear). The labelled oligos are pooled and used to screen 6 x 10^5 plaques from a λZAP-*Caenorhabditis elegans* cDNA library (obtained from Clontech) by published methods [T.N. Davis and J. Thorner *Meth. Enzymol.* 139, 246-262 (1987)]. Hybridization is at 42°C for 16 hours. Positive clones are plaque-purified by twice repeating the hybridization screening process. Plasmid DNAs, excised from phage DNA according to manufacturers instructions, are subjected to automated DNA sequence analysis performed by high temperature cycle sequencing (Applied Biosystems). Geneworks DNA sequence analysis software (Intelligenetics) is used to align raw DNA sequence data and to identify open reading frames.

Example 5

Comparison of the putative proteins encoded by DmORF1 and F22b7.7. Predicted amino acid sequences of DmORF1 and

F22b7.7 are aligned and displayed in FIGURE 4 [SEQ ID NOS:37 and 38]. Only limited overall amino acid homology is exhibited by these two proteins with regions of greatest homology existing in the pore forming H2-1 and H2-2 domains. FIGURE 5A shows a comparison of the pore forming domains of DmORF1 and F22b7.7 with those of the known *Drosophila melanogaster* potassium channel and inward rectifier sequences [SEQ ID NOS:7 through 21]. Amino acid identities greater than 50 % are observed with all potassium channel sequences. FIGURE 5B shows hydropathy plot analysis of DmORF1 and F22b7.7. The two proteins, which show remarkable topological similiarity through their length, are predicted to be composed of four membrane-spanning hydrophobic domains (M1-M4), and two pore forming H2 domains. These data suggest the predicted topology shown in FIGURE 6. Both proteins are predicted to span the membrane four times with amino and carboxyl termini residing within the cell. This topology places the single amino-terminal asparagine-linked glycosylation site and H2 domains on the cell exterior permitting permeation of the membrane by the pore forming domains from the outside, an absolute requirement for the formation of a functional potassium channel.

Example 6

Functional expression of a rat atrial delayed rectifier potassium channel in yeast. CY162 transformants containing plasmids pKAT1, which encodes a plant inward rectifier potassium channel, pRATRAK, which encodes a rat atrial delayed rectifier potassium channel, pDmORF1, and control plasmid pYES are cultured on arginine-phosphate-dextrose agar medium lacking ura medium [A. Rodriguez-Navarro and J. Ramos, *J. Bacteriol.* 159, 940-945, (1984)] containing various KCl c ncentrati ns (FIGURE 7). Strains containing pKAT1, pRATRAK, and pDmORF1 all supp rt the growth of CY162

on medium containing a low concentration of potassium, while pYES2 containing CY162 cells only grow on medium containing a high potassium concentration, indicating that heterologous potassium channels of several different types function to provide high affinity potassium uptake.

pRATRAK is constructed by modifying the protein-coding sequences of RATRAK to add 5' HindIII and 3' XbaI sites using PCR. In addition, four A residues are added to the sequences immediately 5' proximal to the initiator ATG to provide a good yeast translational initiation site. The modified fragment is cloned into the HindIII and XbaI sites in the yeast expression vector pYES2 (Invitrogen), forming pRATRAK.

Example 7

Bioassay of functional expression of heterologous potassium channels

Yeast strains dependent on heterologous potassium channels for growth should be sensitive to non-specific potassium channel blocking compounds. To test the potassium channel blocking properties of several compounds, a convenient agar plate bioassay is employed. Strains containing pKAT1, pRATRAK, pDmORF1, and pYES2 are plated in arginine-phosphate-dextrose agar medium lacking ura and containing various amounts of potassium chloride. Arginine-phosphate-dextrose medium is used to avoid interference from potassium and ammonium ions present in standard synthetic yeast culture medium. Sterile filter disks were placed on the surface of the agar and saturated with potassium channel blocking ions CsCl, $BaCl_2$, and TEA. The growth of heterologous potassium channel containing strains is inhibited by potassium channel bl cking ions, in a channel dependent manner. DmORF1-dependent growth is blocked by

$BaCl_2$ but not by CsCl or TEA. KAT-dependent growth is blocked by $BaCl_2$, CsCl and TEA. RATRAK-dependent growth is blocked by $BaCl_2$, CsCl and TEA to a much greater extent than pKAT1, reflecting in part a slower growth rate of pRATRAK-containing cells. These observations confirm that these channels support the growth of the mutant yeast cells and demonstrate the efficacy of the yeast bioassay for screening for compounds that block potassium channel function. The control pYES-containing strain grows only around applied KCl and RbCl, a congener of KCl.

Example 8

Identification of compounds that alter potassium channel activity

Yeast strains made capable of growing on medium containing low potassium concentration by expression of heterologous potassium channels are used to screen libraries of chemical compounds of diverse structure for those that interfere with channel function. CY162 cells containing pKAT1, pRATRAK, pDmORF1, pCeORF1, and pYES2-*TRK1* (10^4/ml) are plated in 200 ml of arginine-phosphate-dextrose agar medium lacking ura and containing 0.2 mM potassium chloride in 500 cm^2 plates. The CY162 cells bearing pYES2-*TRK1* are included in the assay as a control to identify compounds that have non-specific effects on the yeast strain and are therefore not specifically active against the heterologous potassium channels. Samples of chemical compounds of diverse structure (2 µl of 10 mg/ml solution in DMSO) are applied to the surface of the hardened agar medium in a 24 x 24 array. The plates are incubated f r 2 days at 30°C during which time the applied compounds radially diffuse into the agar medium. The effects of applied compounds on strains bearing heterologous potassium channel

genes are compared to the pYES2-*TRK1* bearing strain. Compounds that cause a zone of growth inhibition around the point of application that is larger on plates containing cells bearing the heterologous potassium channels than that observed around the pYES2-*TRK1* bearing strains are considered selective potassium channel blockers. Compounds that induce a zone of enhanced growth around the point of application that is larger on plates containing cells bearing the heterologous potassium channels than that observed around the pYES2-*TRK1* bearing strains are considered selective potassium channel openers.

Example 9

DmORF1-induced currents in *X. laevis* oocytes assayed by two-electrode voltage clamp

DNA sequence analysis of the pDmORF insert strongly suggest that the protein encoded by the single long ORF possesses properties in common with known potassium channels. To test this hypothesis, the electrophysiological properties of the putative potassium channel encoded by DmORF was examined by expression in *X. laevis* oocytes. Currents were measured by two-electrode whole-cell voltage clamp. DNA sequences encoding the open reading frame of DmORF1 were amplified by polymerase chain reaction (PCR) using the following oligonucleotides:

MPO23: ATAAAGCTTAAAAATGTCGCCGAATCGATGGAT [SEQ ID NO:22]

MPO24: AGCTCTAGACCTCCATCTGGAAGCCCATGT [SEQ ID NO:23]

The full length PCR product was cloned into corresponding sites in pSP64 poly A (Promega), forming pMP147. Template DNA was linearized with EcoRI and RNA transcribed using the Message Machine (Ambion) *in vitro* transcription kit according to manufacturers instructions. A sample of the RNA was resolved in a MOPS-acetate-formaldehyde agarose gel

and RNA content was estimated by ethidium bromide staining. The remainder was stored on dry ice. *X. laevis* oocytes were isolated and injected with 50 nl of sterile TE containing 5-20 ng transcript according to published procedures. After three days, whole oocyte currents were recorded using a two-electrode voltage clamp. Electrodes contained 3M KCl and had resistances of 0.3-1.0 MΩ. Recordings were performed with constant perfusion at room temperature in the presence of either low (10 mM) or high (90 mM) potassium chloride. Two electrode voltage clamp analysis of the DmORF1 gene product expressed in *X. laevis* oocytes demonstrates properties of a voltage- and potassium-dependent potassium channel. At low potassium concentrations, DmORF1 exhibited outward current at depolarizing potentials. At high potassium concentration, DmORF1 exhibits both inward and outward currents. The DmORF1 channel displays a high preference for potassium and shows cation selectivity in the rank order K>Rb>NH_4>Cs>Na>Li. Potassium currents were greatly attenuated by $BaCl_2$.

Example 10

Developmental regulation of DmORF1 expression in *D. melanogaster* determined by northern blotting analysis

Isolation of pDmORF1 from a *D. melanogaster* expression library strongly suggests that the insert contained within originated in mRNA from that species. Detailed understanding of the developmental regulation of DmORF1 expression should aid in determining strategies for use of DmORF1 as a target for novel insecticides. To characterize DmORF1 expression, northern blotting analysis of poly A RNA from various stages of the *D. melanogaster* life cycle was carried out.

D. melanogaster poly A^+ RNA from embryo, larvae and

adult forms (Invitrogen, 5 µg) was resolved in a MOPS-acetate-formaldehyde agarose gel according to standard procedures. The gel was stained with ethidium bromide and photographed to mark the positions of 18 S and 28 S ribosomal RNAs used as molecular weight markers. RNA was transferred by capillary action to nitrocellulose with 10 x SSPE. The blot was air-dried, baked for one hour at 80°C, and prehybridized in 4x SSPE, 1% SDS, 2x Denhardt's, 0.1 % single stranded DNA at 68 °C for 2 hours.

A 2.4 kb *Xho*I fragment of DmORF1 was isolated from pDmORF1 and labeled with α-^{32}P dCTP using the Ready-to-Go kit (Pharmacia) according to manufacturers instructions. The probe was denatured by heating to 100°C for 5 minutes followed by quenching in an ice water bath. The probe was added to the prehybridization solution and hybridization continued for 24 hours at 68 °C.

The blot was washed briefly with 2x SSPE, 0.1% SDS at room temperature followed by 0.5 x SSPE, 0.1 % SDS at 65 °C for 2 hours. The blot was air-dried and exposed to Reflection X-ray film (NEN) using an intensifying screen at -70 °C for 48 hours.

Northern blotting analysis indicates that the DmORF1 probe hybridizes to an mRNA species of approximately 2.8 kb isolated from *D. melanogaster* embryo, larvae, and adult forms. The length of the DmORF1 mRNA corresponds well with the length of the predicted ORF. Thus, the DmORF is expressed at all developmental stages in the life cycle of *D. melanogaster*.

Example 11

Expression of the DmORF1 gene product *in vitro*.

DNA sequence analysis of the pDmORF1 insert reveals a single long ORF with conserved amino acid sequence domains

in common with known potassium channels. The DNA sequence predicts an ORF sufficient to encode a protein of 618 amino acid in length. The DmORF1 polypeptide contains four segments of at least 20 hydrophobic amino acids in length suggesting that the segments span the plasma membrane. In addition, the DmORF1 protein sequence contains a putative N-linked glycosylation site (Asn-Thr-Thr) at amino acids 58-60. To confirm that a protein of the predicted size of DmORF is expressed from the insert in pDmORF1 and to test the proposition that DmORF1 is glycosylated, pDmORF1 was used as template to drive coupled *in vitro* transcription/translation.

Plasmid pMP147 was used as template to produce ^{35}S-labeled DmORF gene product *in vitro* using a TnT coupled transcription-translation kit (Promega) according to manufacturers instructions. Glycosylation of the nascent DmORF1 polypeptide was accomplished by addition of canine pancreatic microsomes (Promega) to the transcription-translation reaction. Samples of glycosylated DmORF protein were treated with endoglycosidase H to remove added carbohydrate moieties. Aliquots were precipitated with TCA and collected on GF/C filters, washed with ethanol, dried and counted. Equivalent cpm's were resolved by SDS-PAGE. The gel was impregnated with soluble fluor Amplify (Amersham) and dried onto Whatman 3MM paper. The dried gel was exposed to Reflection X-ray film at room temperature.

Translation of the DmORF1 gene product *in vitro* produced a polypeptide of 68 kDa, consistent with the predicted molecular weight of the ORF. Translation of DmORF1 in the presence of canine pancreatic microsomes results in synthesis of a protein with reduced electrophoretic mobility, consistent with glycosylation of the nascent p lypeptide. Treatment of glycosylated DmORF with EndoH increased its relative mobility as expected upon

removal of carbohydrate moieties. Thus, the pDmORF1 insert is capable of directing the expression of a glycoprotein with the expected molecular weight. EndoH treatment removes carbohydrate residues consistent with the sugar added through N-linked glycosylation.

Example 12

High-affinity K^+ uptake and selectivity of DmORF1 expressed in yeast.

Expression of DmORF permits CY162 cells to grow on medium containing a low concentration of potassium, implying that DmORF1 supplies high affinity potassium uptake capacity. To characterize the potassium uptake properties of CY162 cells containing DmORF1, ^{86}Rb uptake studies were performed. Examination of the uptake of this potassium congener revealed important aspects of potassium uptake by DmORF1.

Yeast strains containing heterologous potassium-expression plasmids CY162-DmORF1, CY162-pKAT and the control strain CY162-pYES2 (Clontech) were cultured overnight in SC Gal-ura containing 0.1 M KCl. The cells were harvested, washed with sterile doubled distilled water and starved for K^+ for 6 hours in Ca-MES buffer. Cells were washed again and distributed to culture tubes (10^8 cells/tube) containing ^{86}RbCl in Ca-MES buffer. The tubes were incubated at room temperature, samples filtered at various time intervals and counted. ^{86}Rb uptake into cells was displayed. For Double Reciprocal analysis, ^{86}Rb was held constant and barium ions varied to determine Ki values.

Th high-affinity potassium uptake capacity encoded by DmORF1 permits high-affinity uptake f the potassium congener, ^{86}Rb, as well. Barium inhibited ^{86}Rb uptake with a

Ki of µM as demonstrated in Double Reciprocal analysis. No high affinity ^{86}Rb uptake is observed in control CY162-pYES2 cells and ^{86}Rb uptake into CY162-pKAT cells is consistent with its published properties.

Example 13

Expression of *Drosophila melanogaster* potassium channels in yeast.

Voltage-gated potassium channel diversity in the fruitfly *Drosophila melanogaster* is encoded in large part by six genes, Shaker, Shab, Shal, Shaw, Eag, and Slo. Expression of these potassium channels in yeast will permit their introduction into screening assays for novel insecticidal compounds and facilitate characterization of their ion channel properties and sensitivity to compounds with activating and inhibitory properties.

DNA sequences encoding *Drosophila melanogaster* potassium channels were amplified by PCR using synthetic oligonucleotides that add 5' HindIII or Kpn I, sites and 3' XbaI, SphI, or XhoI sites:

Shaker 5':AAAAAGCTTAAAATGGCACACATCACG [SEQ ID NO:24]
Shaker 3':AAACTCGAGTCATACCTGTGGACT [SEQ ID NO:25]

Shab 5':AAAAAGCTTAAAATGGTCGGGCAATTG [SEQ ID NO:26]
Shab 3':AAAAGCATGCTCATCTGGATGGGCA [SEQ ID NO:27]

Shal 5':AAAAAGCTTAAAATGGCCTCGGTCGCC [SEQ ID NO:28]
Shal 3':TTTTCTAGACTACATCGTTGTCTT [SEQ ID NO:29]

Shaw 5':AAAAAGCTTAAAATGAATCTGATCAAC [SEQ ID NO:30]
Shaw 3':AAATCTAGATTAGTCGAAACTGAA [SEQ ID NO:31]

Eag 5':AAAAAGCTTAAAATGCCTGGCGGA [SEQ ID NO:32]
Eag 3':AAATCTAGAGGCTACAGGAAGTCC [SEQ ID NO:33]

Slo 5':GGGGGTACCAAAATGTCGGGGTGTGAT [SEQ ID NO:34]
Slo 3':TTTTTCTAGATCAAGAGTTATCATC [SEQ ID NO:35]

Plasmids used as templates for the PCR reactions were: pBSc-DShakerH37, pBSc-dShab11, pBSc-dShal2+$(A)_{36}$, pBScMXT-dShaw [A. Wei, M. Covarrubias, A. Butler, K. Baker, M. Pak, L. Salkoff, *Science* 248, 599-603 (1990), provided by L. Salkoff], pBScMXT-slo,v4 [N.S. Atkinson, G.A. Robertson, B. Ganetzky, *Science* 253,551-555, (1991), provided by L. Salkoff], and pBIMCH20 Eag [CH20] [J. Warmke, R. Drysdale, B. Ganetzky, *Science* 252, 1560-1564 (1991), A. Bruggemann, L.A. Pardo, W. Stuhmer, O. Pongs, *Nature* 365, 445-448 (1993), provided by B. Ganetzky].

Amplified fragments were digested with the appropriate restriction endonucleases, purified using GeneClean (Bio 101), and ligated into corresponding sites in pYES2 (Invitrogen). CY162 cells were transformed with assembled *Drosophila melanogaster* potassium channel expression plasmids by the LiCl method and plated on SCD-ura containing 0.1M KCl agar medium. Selected transformants were tested for growth on arginine-phosphate-galactose (2 %)/sucrose (0.2 %)-ura agar medium containing 1-5 mM KCl. CY162 cells containing pKAT1 or pDmORF1 were cultured as positive controls and CY162 cells containing pYES2 were grown to provide a negative control.

CY162 cells bearing *Drosophila melanogaster* potassium channel expression plasmids survive under conditions in which growth is dependent on functional potassium channel expression. At potassium i n concentrati ns between 1-3 mM, negative control CY162 cells containing pYES2 grow poorly. Expression of the *Drosophila melanogaster* potassium channels

Shal, Shaw and Eag substantially improve growth of CY162. These results are consistent with the *Drosophila melanogaster* potassium channels providing high-affinity potassium uptake capacity. This capacity is apparently sufficient to replace the native high-affinity potassium transport capacity encoded by TRK1 which is lacking in CY162 (*trk1 trk2*) cells.

Example 14

Cloning of a novel *C. elegans* sequence with homology to potassium channels.

In order to expand the applicability of this technology to discover compounds with novel anhelmenthic activity, CY162 cells were transformed with a pYES2-based yeast expression library constructed using cDNA synthesized from *C. elegans* mRNA (Invitrogen). Plasmid DNA isolated from yeast cells that survived the selection scheme described in EXAMPLE 1 were subjected to automated DNA sequence analysis performed by high temperature cycle sequencing (Applied Biosystems). Geneworks DNA sequence analysis software (Intelligenetics) is used to align raw DNA sequence information and to identify open reading frames. The DNA sequence of the 1.4 kb insert in pCORK is displayed in FIGURE 9A and 9B. The 5' untranslated sequences of the cDNA are present in this construct. A single long open reading frame sufficient to encode a protein of 434 amino acids (predicted MW 48 kDa) is predicted in pCORK [SEQ ID NO:38]. A consensus polyadenylation site, AATAAA, occurs at position 1359-1364 in 3' untranslated sequences and is followed by a tract of 15 consecutive A residues. The CORK ORF contains structural features that resemble p re forming H5 domains found in potassium channels. Two putative pore forming H5 domains (residues 76-39 and 150-162) contain the G-Y/F-G

tripeptide motif required for potassium selectivity [L. Heginbotham, T. Abramson, R. MacKinnon, Science 258, 1152-1155, (1992)].

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: American Cyanamid Company

(ii) TITLE OF INVENTION: Genes Encoding a Novel Family of Potassium Channels

(iii) NUMBER OF SEQUENCES: 38

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: American Cyanamid Company
(B) STREET: One Cyanamid Plaza
(C) CITY: Wayne
(D) STATE: New Jersey
(E) COUNTRY: USA
(F) ZIP: 07470-8426

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Matthews, Gale F.
(B) REGISTRATION NUMBER: 32,369
(C) REFERENCE/DOCKET NUMBER: 32,421-01 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 201-660-6329
(B) TELEFAX: 201-660-7160

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2441 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 190..2043

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ACGCGATCGC CGCGAGTGTA TATTTTTTTT TTAGCTCAGT CTTCAGTGTT TCGCGATTCT       60

CTTTAAAAGA AAAAAAAAAT AATAAGTCAA AACTACAAAC CACACAGCGA AAGGCGAAAG      120

CAACGGTTCC TGCGAGTGTT TATTTTTTTT TTCAACAATT TTTGATCGTA GTGCGACAAT      180

CCGTCGAGC ATG TCG CCG AAT CGA TGG ATC CTG CTG CTC ATC TTC TAC          228
          Met Ser Pro Asn Arg Trp Ile Leu Leu Leu Ile Phe Tyr
            1               5                  10

ATA TCC TAC CTG ATG TTC GGG GCG GCA ATC TAT TAC CAT ATT GAG CAC        276
Ile Ser Tyr Leu Met Phe Gly Ala Ala Ile Tyr Tyr His Ile Glu His
```

```
   15                  20                  25

GGC GAG GAG AAG ATA TCG CGC GCC GAA CAG CGC AAG GCG CAA ATT GCA      324
Gly Glu Glu Lys Ile Ser Arg Ala Glu Gln Arg Lys Ala Gln Ile Ala
 30                 35                  40                  45

ATC AAC GAA TAT CTG CTG GAG GAG CTG GGC GAC AAG AAT ACG ACC ACA      372
Ile Asn Glu Tyr Leu Leu Glu Glu Leu Gly Asp Lys Asn Thr Thr Thr
                 50                  55                  60

CAG GAT GAG ATT CTT CAA CGG ATC TCG GAT TAC TGT GAC AAA CCG GTT      420
Gln Asp Glu Ile Leu Gln Arg Ile Ser Asp Tyr Cys Asp Lys Pro Val
             65                  70                  75

ACA TTG CCG CCG ACA TAT GAT GAT ACG CCC TAC ACG TGG ACC TTC TAC      468
Thr Leu Pro Pro Thr Tyr Asp Asp Thr Pro Tyr Thr Trp Thr Phe Tyr
         80                  85                  90

CAT GCC TTC TTC TTC GCC TTC ACC GTT TGC TCC ACG GTG GGA TAT GGG      516
His Ala Phe Phe Phe Ala Phe Thr Val Cys Ser Thr Val Gly Tyr Gly
     95                 100                 105

AAT ATA TCG CCA ACC ACC TTC GCC GGA CGG ATG ATC ATG ATC GCG TAT      564
Asn Ile Ser Pro Thr Thr Phe Ala Gly Arg Met Ile Met Ile Ala Tyr
110                 115                 120                 125

TCG GTG ATT GGC ATC CCC GTC AAT GGT ATC CTC TTT GCC GGC CTC GGC      612
Ser Val Ile Gly Ile Pro Val Asn Gly Ile Leu Phe Ala Gly Leu Gly
                130                 135                 140

GAA TAC TTT GGA CGT ACG TTT GAA GCG ATC TAC AGA CGC TAC AAA AAG      660
Glu Tyr Phe Gly Arg Thr Phe Glu Ala Ile Tyr Arg Arg Tyr Lys Lys
            145                 150                 155

TAC AAG ATG TCC ACG GAT ATG CAC TAT GTC CCG CCG CAG CTG GGA TTG      708
Tyr Lys Met Ser Thr Asp Met His Tyr Val Pro Pro Gln Leu Gly Leu
        160                 165                 170

ATC ACC ACG GTG GTG ATT GCC CTG ATT CCG GGA ATA GCT CTC TTC CTG      756
Ile Thr Thr Val Val Ile Ala Leu Ile Pro Gly Ile Ala Leu Phe Leu
    175                 180                 185

GTG CTG CCC TGC GTG GGT GTT CAC CTA CTT CGA GAA CTG GGC CTA TCT      804
Val Leu Pro Cys Val Gly Val His Leu Leu Arg Glu Leu Gly Leu Ser
190                 195                 200                 205

TCC ATC TCG CTG TAC TAC AGC TAT GTG ACC ACC ACA ACA ATT GGA TTC      852
Ser Ile Ser Leu Tyr Tyr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe
                210                 215                 220

GGT GAC TAT GTG CCC ACA TTT GGA GCC AAC CAG CCC AAG GAG TTC GGC      900
Gly Asp Tyr Val Pro Thr Phe Gly Ala Asn Gln Pro Lys Glu Phe Gly
            225                 230                 235

GGC TGG TTC GTG GTC TAT CAG ATC TTT GTG ATC GTG TGG TTC ATC TTC      948
Gly Trp Phe Val Val Tyr Gln Ile Phe Val Ile Val Trp Phe Ile Phe
        240                 245                 250

TCG CTG GGA TAT CTT GTG ATG ATC ATG ACA TTT ATC ACT CGG GGC CTC      996
Ser Leu Gly Tyr Leu Val Met Ile Met Thr Phe Ile Thr Arg Gly Leu
    255                 260                 265

CAG AGC AAG AAG CTG GCA TAC CTG GAG CAG CAG TTG TCC TCC AAC CTG     1044
Gln Ser Lys Lys Leu Ala Tyr Leu Glu Gln Gln Leu Ser Ser Asn Leu
270                 275                 280                 285

AAG GCC ACA CAG AAT CGC ATC TGG TCT GGC GTC ACC AAG GAT GTG GGC     1092
Lys Ala Thr Gln Asn Arg Ile Trp Ser Gly Val Thr Lys Asp Val Gly
                290                 295                 300
```

TAC CTC CGG CGA ATG CTC AAC GAG CTG TAC ATC CTC AAA GTG AAG CCT 1140
Tyr Leu Arg Arg Met Leu Asn Glu Leu Tyr Ile Leu Lys Val Lys Pro
305 310 315

GTG TAC ACC GAT GTA GAT ATC GCC TAC ACA CTG CCA CGT TCC AAT TCG 1188
Val Tyr Thr Asp Val Asp Ile Ala Tyr Thr Leu Pro Arg Ser Asn Ser
320 325 330

TGT CCG GAT CTG AGC ATG TAC CGC GTG GAG CCG GCT CCC ATT CCC AGC 1236
Cys Pro Asp Leu Ser Met Tyr Arg Val Glu Pro Ala Pro Ile Pro Ser
335 340 345

CGG AAG AGG GCA TTC TCC GTG TGC GCC GAC ATG GTT GGC GCC CAA AGG 1284
Arg Lys Arg Ala Phe Ser Val Cys Ala Asp Met Val Gly Ala Gln Arg
350 355 360 365

GAG GCG GGC ATG GTA CAC GCC AAT TCC GAT ACG GAT CTA ACC AAA CTG 1332
Glu Ala Gly Met Val His Ala Asn Ser Asp Thr Asp Leu Thr Lys Leu
370 375 380

GAT CGC GAG AAG ACA TTC GAG ACG GCG GAG GCG TAC CAC CAG ACC ACC 1380
Asp Arg Glu Lys Thr Phe Glu Thr Ala Glu Ala Tyr His Gln Thr Thr
385 390 395

GAT TTG CTG GCC AAG GTG GTC AAC GCA CTG GCC ACG GTG AAG CCA CCG 1428
Asp Leu Leu Ala Lys Val Val Asn Ala Leu Ala Thr Val Lys Pro Pro
400 405 410

CCG GCG GAA CAG GAA GAT GCG GCT CTC TAT GGT GGC TAT CAT GGC TTC 1476
Pro Ala Glu Gln Glu Asp Ala Ala Leu Tyr Gly Gly Tyr His Gly Phe
415 420 425

TCC GAC TCC CAG ATC CTG GCC AGC GAA TGG TCG TTC TCG ACG GTC AAC 1524
Ser Asp Ser Gln Ile Leu Ala Ser Glu Trp Ser Phe Ser Thr Val Asn
430 435 440 445

GAG TTC ACA TCA CCG CGA CGT CCA AGA GCA CGT GCC TGC TCC GAT TTC 1572
Glu Phe Thr Ser Pro Arg Arg Pro Arg Ala Arg Ala Cys Ser Asp Phe
450 455 460

AAT CTG GAG GCA CCT CGC TGG CAG AGC GAG AGG CCA CTG CGT TCG AGC 1620
Asn Leu Glu Ala Pro Arg Trp Gln Ser Glu Arg Pro Leu Arg Ser Ser
465 470 475

CAC AAC GAA TGG ACA TGG AGC GGC GAC AAC CAG CAG ATC CAG GAG GCA 1668
His Asn Glu Trp Thr Trp Ser Gly Asp Asn Gln Gln Ile Gln Glu Ala
480 485 490

TTC AAC CAG CGC TAC AAG GGA CAG CAG CGT GCC AAC GGA GCA GCC AAC 1716
Phe Asn Gln Arg Tyr Lys Gly Gln Gln Arg Ala Asn Gly Ala Ala Asn
495 500 505

TCG ACC ATG GTC CAT CTG GAG CCG GAT GCT TTG GAG GAG CAG CTG AGA 1764
Ser Thr Met Val His Leu Glu Pro Asp Ala Leu Glu Glu Gln Leu Arg
510 515 520 525

AAC AAT CAC CGG GTG CCG GTC GCG TCA AGA AGT TCT CCA TGC CGG ATG 1812
Asn Asn His Arg Val Pro Val Ala Ser Arg Ser Ser Pro Cys Arg Met
530 535 540

GTC TGC GAC GTC TGT TTC CCT TCC AGA AGA AGC ACC CCT CGC AGG ATC 1860
Val Cys Asp Val Cys Phe Pro Ser Arg Arg Ser Thr Pro Arg Arg Ile
545 550 555

TGG AGC GCA AGT TGT CCG TGG TCT CGG TAC CCG AGG GTG TCA TCT CGC 1908
Trp Ser Ala Ser Cys Pr Trp Ser Arg Tyr Pro Arg Val Ser Ser Arg
560 565 570

AGG AAG CCA GAT CCC CGC TGG ACT ACT ACA TCA ACA CGG TCA CGG CGG 1956

Arg Lys Pro Asp Pro Arg Trp Thr Thr Thr Ser Thr Arg Ser Arg Arg
575 580 585

CCT CCA GTC AAT CCT ATT TGC GCA ACG GAC GCG GTC CGC CAC CGC CCT 2004
Pro Pro Val Asn Pro Ile Cys Ala Thr Asp Ala Val Arg His Arg Pro
590 595 600 605

TCG AAT CGA ATG GCA GCT TGG CCA GCG GCG GCG GCG GGC TAACGAACAT 2053
Ser Asn Arg Met Ala Ala Trp Pro Ala Ala Ala Ala Gly
610 615

GGGCTTCCAG ATGGAGGATG GAGCAACCCC GCCATCGGCA TTGGGCGGTG GAGCCTATCA 2113

ACGCAAGGCG GCTGCTGGCA AGCGCCGACG CGAGAGCATC TACACCCAGA ATCAAGCCCC 2173

ATCCGCTCGC CGGGGCAGCA TGTATCCGCC GACCGCGCAC GCCTTGGCCC AGATGCAGAT 2233

GCGACGCGGC AGCTTGGCAA CCAGTGGCTC TGGATCGGCG GCCATGGCGG CAGTGGCCGC 2293

GCGTCGTGGC AGCCTCTTCC CAGCTACAGC ATCGGCATCA TCGCTGACCT CTGCTCCGCG 2353

CCGAAGCAGC ATATTCTCGG TTACCTCCGA AAAGGATATG AATGTGCTGG AGCAGACGAC 2413

CATTGCGGAT CTGATTCGTG CGCTCGAG 2441

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 618 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Pro Asn Arg Trp Ile Leu Leu Leu Ile Phe Tyr Ile Ser Tyr
1 5 10 15

Leu Met Phe Gly Ala Ala Ile Tyr Tyr His Ile Glu His Gly Glu Glu
20 25 30

Lys Ile Ser Arg Ala Glu Gln Arg Lys Ala Gln Ile Ala Ile Asn Glu
35 40 45

Tyr Leu Leu Glu Glu Leu Gly Asp Lys Asn Thr Thr Thr Gln Asp Glu
50 55 60

Ile Leu Gln Arg Ile Ser Asp Tyr Cys Asp Lys Pro Val Thr Leu Pro
65 70 75 80

Pro Thr Tyr Asp Asp Thr Pro Tyr Thr Trp Thr Phe Tyr His Ala Phe
85 90 95

Phe Phe Ala Phe Thr Val Cys Ser Thr Val Gly Tyr Gly Asn Ile Ser
100 105 110

Pro Thr Thr Phe Ala Gly Arg Met Ile Met Ile Ala Tyr Ser Val Ile
115 120 125

Gly Ile Pro Val Asn Gly Ile Leu Phe Ala Gly Leu Gly Glu Tyr Phe
130 135 140

Gly Arg Thr Phe Glu Ala Ile Tyr Arg Arg Tyr Lys Lys Tyr Lys Met
145 150 155 160

Ser Thr Asp Met His Tyr Val Pr Pro Gln Leu Gly Leu Ile Thr Thr
165 170 175

```
Val Val Ile Ala Leu Ile Pro Gly Ile Ala Leu Phe Leu Val Leu Pro
            180                 185                 190

Cys Val Gly Val His Leu Leu Arg Glu Leu Gly Leu Ser Ser Ile Ser
        195                 200                 205

Leu Tyr Tyr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe Gly Asp Tyr
    210                 215                 220

Val Pro Thr Phe Gly Ala Asn Gln Pro Lys Glu Phe Gly Gly Trp Phe
225                 230                 235                     240

Val Val Tyr Gln Ile Phe Val Ile Val Trp Phe Ile Phe Ser Leu Gly
                245                 250                 255

Tyr Leu Val Met Ile Met Thr Phe Ile Thr Arg Gly Leu Gln Ser Lys
            260                 265                 270

Lys Leu Ala Tyr Leu Glu Gln Gln Leu Ser Ser Asn Leu Lys Ala Thr
        275                 280                 285

Gln Asn Arg Ile Trp Ser Gly Val Thr Lys Asp Val Gly Tyr Leu Arg
    290                 295                 300

Arg Met Leu Asn Glu Leu Tyr Ile Leu Lys Val Lys Pro Val Tyr Thr
305                 310                 315                     320

Asp Val Asp Ile Ala Tyr Thr Leu Pro Arg Ser Asn Ser Cys Pro Asp
                325                 330                 335

Leu Ser Met Tyr Arg Val Glu Pro Ala Pro Ile Pro Ser Arg Lys Arg
            340                 345                 350

Ala Phe Ser Val Cys Ala Asp Met Val Gly Ala Gln Arg Glu Ala Gly
        355                 360                 365

Met Val His Ala Asn Ser Asp Thr Asp Leu Thr Lys Leu Asp Arg Glu
    370                 375                 380

Lys Thr Phe Glu Thr Ala Glu Ala Tyr His Gln Thr Thr Asp Leu Leu
385                 390                 395                     400

Ala Lys Val Val Asn Ala Leu Ala Thr Val Lys Pro Pro Pro Ala Glu
                405                 410                 415

Gln Glu Asp Ala Ala Leu Tyr Gly Gly Tyr His Gly Phe Ser Asp Ser
            420                 425                 430

Gln Ile Leu Ala Ser Glu Trp Ser Phe Ser Thr Val Asn Glu Phe Thr
        435                 440                 445

Ser Pro Arg Arg Pro Arg Ala Arg Ala Cys Ser Asp Phe Asn Leu Glu
    450                 455                 460

Ala Pro Arg Trp Gln Ser Glu Arg Pro Leu Arg Ser Ser His Asn Glu
465                 470                 475                     480

Trp Thr Trp Ser Gly Asp Asn Gln Gln Ile Gln Glu Ala Phe Asn Gln
                485                 490                 495

Arg Tyr Lys Gly Gln Gln Arg Ala Asn Gly Ala Ala Asn Ser Thr Met
            500                 505                 510

Val His Leu Glu Pro Asp Ala Leu Glu Glu Gln Leu Arg Asn Asn His
        515                 520                 525

Arg Val Pro Val Ala Ser Arg Ser Ser Pro Cys Arg Met Val Cys Asp
    530                 535                 540
```

```
Val Cys Phe Pro Ser Arg Arg Ser Thr Pro Arg Arg Ile Trp Ser Ala
545                 550                 555                 560

Ser Cys Pro Trp Ser Arg Tyr Pro Arg Val Ser Ser Arg Arg Lys Pro
                565                 570                 575

Asp Pro Arg Trp Thr Thr Thr Ser Thr Arg Ser Arg Arg Pro Pro Val
            580                 585                 590

Asn Pro Ile Cys Ala Thr Asp Ala Val Arg His Arg Pro Ser Asn Arg
        595                 600                 605

Met Ala Ala Trp Pro Ala Ala Ala Ala Gly
    610                 615
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1011 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1008

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATG TCC GAT CAG CTG TTT GTC GCA TTT GAG AAG TAT TTC TTG ACG AGT      48
Met Ser Asp Gln Leu Phe Val Ala Phe Glu Lys Tyr Phe Leu Thr Ser
  1               5                  10                  15

AAC GAG GTC AAG AAG AAT GCA GCA ACG GAG ACA TGG ACA TTT TCA TCG      96
Asn Glu Val Lys Lys Asn Ala Ala Thr Glu Thr Trp Thr Phe Ser Ser
             20                  25                  30

TCC ATT TTC TTT GCC GTA ACC GTC GTC ACT ACC ATC GGA TAC GGT AAT     144
Ser Ile Phe Phe Ala Val Thr Val Val Thr Thr Ile Gly Tyr Gly Asn
         35                  40                  45

CCA GTT CCA GTG ACA AAC ATT GGA CGG ATA TGG TGT ATA TTG TTC TCC     192
Pro Val Pro Val Thr Asn Ile Gly Arg Ile Trp Cys Ile Leu Phe Ser
     50                  55                  60

TTG CTT GGA ATA CCT CTA ACA CTG GTT ACC ATC GCT GAC TTG GCA GGT     240
Leu Leu Gly Ile Pro Leu Thr Leu Val Thr Ile Ala Asp Leu Ala Gly
 65                  70                  75                  80

AAA TTC CTA TCT GAA CAT CTT GTT TGG TTG TAT GGA AAC TAT TTG AAA     288
Lys Phe Leu Ser Glu His Leu Val Trp Leu Tyr Gly Asn Tyr Leu Lys
                 85                  90                  95

TTA AAA TAT CTC ATA TTG TCA CGA CAT CGA AAA GAA CGG AGA GAG CAC     336
Leu Lys Tyr Leu Ile Leu Ser Arg His Arg Lys Glu Arg Arg Glu His
            100                 105                 110

GTT TGT GAG CAC TGT CAC AGT CAT GGA ATG GGG CAT GAT ATG AAT ATC     384
Val Cys Glu His Cys His Ser His Gly Met Gly His Asp Met Asn Ile
        115                 120                 125

GAG GAG AAA AGA ATT CCT GCA TTC CTG GTA TTA GCT ATT CTG ATA GTA     432
Glu Glu Lys Arg Ile Pro Ala Phe Leu Val Leu Ala Ile Leu Ile Val
    130                 135                 140

TAT ACA GCG TTT GGC GGT GTC CTA ATG TCA AAA TTA GAG CCG TGG TCT     480
Tyr Thr Ala Phe Gly Gly Val Leu Met Ser Lys Leu Glu Pro Trp Ser
```

```
145                 150                 155                 160

TTC TTC ACT TCA TTC TAC TGG TCC TTC ATT ACA ATG ACT ACT GTC GGG     528
Phe Phe Thr Ser Phe Tyr Trp Ser Phe Ile Thr Met Thr Thr Val Gly
                165                 170                 175

TTT GGC GAC TTG ATG CCC AGA AGG GAC GGA TAC ATG TAT ATC ATA TTG     576
Phe Gly Asp Leu Met Pro Arg Arg Asp Gly Tyr Met Tyr Ile Ile Leu
            180                 185                 190

CTC TAT ATC ATT TTA GGT AAA TTT TCA ATG AAA AAA AAA CAA AAA TTC     624
Leu Tyr Ile Ile Leu Gly Lys Phe Ser Met Lys Lys Lys Gln Lys Phe
        195                 200                 205

AAA ATA TTT TTA GGT CTT GCA ATA ACT ACA ATG TGC ATT GAT TTG GTA     672
Lys Ile Phe Leu Gly Leu Ala Ile Thr Thr Met Cys Ile Asp Leu Val
    210                 215                 220

GGA GTA CAG TAT ATT CGA AAG ATT CAT TAT TTC GGA AGA AAA ATT CAA     720
Gly Val Gln Tyr Ile Arg Lys Ile His Tyr Phe Gly Arg Lys Ile Gln
225                 230                 235                 240

GAC GCT AGA TCT GCA TTG GCG GTT GTA GGA GGA AAG GTA GTC CTT GTA     768
Asp Ala Arg Ser Ala Leu Ala Val Val Gly Gly Lys Val Val Leu Val
                245                 250                 255

TCA GAA CTC TAC GCA AAT TTA ATG CAA AAG CGA GCT CGT AAC ATG TCC     816
Ser Glu Leu Tyr Ala Asn Leu Met Gln Lys Arg Ala Arg Asn Met Ser
            260                 265                 270

CGA GAA GCT TTT ATA GTG GAG AAT CTC TAT GTT TCC AAA CAC ATC ATA     864
Arg Glu Ala Phe Ile Val Glu Asn Leu Tyr Val Ser Lys His Ile Ile
        275                 280                 285

CCA TTC ATA CCA ACT GAT ATC CGA TGT ATT CGA TAT ATT GAT CAA ACT     912
Pro Phe Ile Pro Thr Asp Ile Arg Cys Ile Arg Tyr Ile Asp Gln Thr
    290                 295                 300

GCC GAT GCT GCT ACC ATT TCC ACG TCA TCG TCT GCA ATT GAT ATG CAA     960
Ala Asp Ala Ala Thr Ile Ser Thr Ser Ser Ser Ala Ile Asp Met Gln
305                 310                 315                 320

AGT TGT AGA TTT TGT CAT TCA AGA TAT TCT CTC AAT CGT GCA TTC AAA    1008
Ser Cys Arg Phe Cys His Ser Arg Tyr Ser Leu Asn Arg Ala Phe Lys
                325                 330                 335

TAG                                                                1011
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 336 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Ser Asp Gln Leu Phe Val Ala Phe Glu Lys Tyr Phe Leu Thr Ser
1               5                   10                  15

Asn Glu Val Lys Lys Asn Ala Ala Thr Glu Thr Trp Thr Phe Ser Ser
            20                  25                  30

Ser Il  Phe Phe Ala Val Thr Val Val Thr Thr Ile Gly Tyr Gly Asn
        35                  40                  45
```

Pro Val Pro Val Thr Asn Ile Gly Arg Ile Trp Cys Ile Leu Phe Ser
50 55 60

Leu Leu Gly Ile Pro Leu Thr Leu Val Thr Ile Ala Asp Leu Ala Gly
65 70 75 80

Lys Phe Leu Ser Glu His Leu Val Trp Leu Tyr Gly Asn Tyr Leu Lys
85 90 95

Leu Lys Tyr Leu Ile Leu Ser Arg His Arg Lys Glu Arg Arg Glu His
100 105 110

Val Cys Glu His Cys His Ser His Gly Met Gly His Asp Met Asn Ile
115 120 125

Glu Glu Lys Arg Ile Pro Ala Phe Leu Val Leu Ala Ile Leu Ile Val
130 135 140

Tyr Thr Ala Phe Gly Gly Val Leu Met Ser Lys Leu Glu Pro Trp Ser
145 150 155 160

Phe Phe Thr Ser Phe Tyr Trp Ser Phe Ile Thr Met Thr Thr Val Gly
165 170 175

Phe Gly Asp Leu Met Pro Arg Arg Asp Gly Tyr Met Tyr Ile Ile Leu
180 185 190

Leu Tyr Ile Ile Leu Gly Lys Phe Ser Met Lys Lys Lys Gln Lys Phe
195 200 205

Lys Ile Phe Leu Gly Leu Ala Ile Thr Thr Met Cys Ile Asp Leu Val
210 215 220

Gly Val Gln Tyr Ile Arg Lys Ile His Tyr Phe Gly Arg Lys Ile Gln
225 230 235 240

Asp Ala Arg Ser Ala Leu Ala Val Val Gly Gly Lys Val Val Leu Val
245 250 255

Ser Glu Leu Tyr Ala Asn Leu Met Gln Lys Arg Ala Arg Asn Met Ser
260 265 270

Arg Glu Ala Phe Ile Val Glu Asn Leu Tyr Val Ser Lys His Ile Ile
275 280 285

Pro Phe Ile Pro Thr Asp Ile Arg Cys Ile Arg Tyr Ile Asp Gln Thr
290 295 300

Ala Asp Ala Ala Thr Ile Ser Thr Ser Ser Ser Ala Ile Asp Met Gln
305 310 315 320

Ser Cys Arg Phe Cys His Ser Arg Tyr Ser Leu Asn Arg Ala Phe Lys
325 330 335

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCCATTTTCT TTGCCGTAAC CGTCGTCACT ACCATCGGAT ACGGTAATCC A 51

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCATTCTACT GGTCCTTCAT TACAATGACT ACTGTCGGGT TTGGCGACTT G 51

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ala Phe Leu Phe Ser Ile Glu Thr Gln Thr Thr Ile Gly Tyr Gly Phe
1 5 10 15

Arg Cys Val Thr Asp Glu Cys Pro
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Phe Leu Phe Ser Leu Glu Thr Gln Val Thr Ile Gly Tyr Gly Phe
1 5 10 15

Arg Cys Val Thr Glu Gln Cys Ala
20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ala Phe Leu Phe Phe Ile Glu Thr Glu Ala Thr Ile Gly Tyr Gly Tyr
1 5 10 15

Arg Tyr Il Thr Asp His Cys Pro
20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ala Phe Phe Phe Ala Phe Thr Val Cys Ser Thr Val Gly Tyr Gly Asn
1 5 10 15

Ile Ser Pro Thr Thr Phe Ala Gly
20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ala Phe Trp Trp Ala Val Val Thr Met Thr Thr Val Gly Tyr Gly Asp
1 5 10 15

Met Thr Pro Val Gly Phe Trp Gly
20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ala Phe Trp Tyr Thr Ile Val Thr Met Thr Thr Leu Gly Tyr Gly Asp
1 5 10 15

Met Val Pro Glu Thr Ile Ala Gly
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala Phe Trp Trp Ala Gly Ile Thr Met Thr Thr Val Gly Tyr Gly Asp
1 5 10 15

Ile Cys Pro Thr Thr Ala Leu Gly
20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly Leu Trp Trp Ala Leu Val Thr Met Thr Thr Val Gly Tyr Gly Asp
1 5 10 15

Met Ala Pro Lys Thr Tyr Ile Gly
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Leu Tyr Phe Thr Met Thr Cys Met Thr Ser Val Gly Phe Gly Asn
1 5 10 15

Val Ala Ala Glu Thr Asp Asn Glu
20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Cys Val Tyr Phe Leu Ile Val Thr Met Ser Thr Val Gly Tyr Gly Asp
1 5 10 15

Val Tyr Cys Glu Thr Val Leu Gly
20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ser Leu Tyr Thr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe Gly Asp
1 5 10 15

Tyr Val Pro Thr Phe Gly Ala Asn
20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Phe Phe Phe Ala Phe Thr Val Cys Ser Thr Val Gly Tyr Gly Asn
1 5 10 15

Ile Ser Pro Thr Thr Phe Ala Gly
20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ser Ile Phe Phe Ala Val Thr Val Val Thr Thr Ile Gly Tyr Gly Asn
1 5 10 15

Pro Val Pro Val Thr Asn Thr Gly
20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ser Leu Tyr Thr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe Gly Asp
1 5 10 15

Tyr Val Pro Thr Phe Gly Ala Asn
20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser Phe Tyr Trp Ser Phe Ile Th. Met Thr Thr Val Gly Ph Gly Asp
1 5 10 15

Leu Met Pro Arg Arg Asp Gly Tyr
20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATAAAGCTTA AAAATGTCGC CGAATCGATG GAT 33

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGCTCTAGAC CTCCATCTGG AAGCCCATGT 30

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AAAAAGCTTA AAATGGCACA CATCACG 27

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAACTCGAGT CATACCTGTG GACT 24

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AAAAAGCTTA AAATGGTCGG GCAATTG 27

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AAAAGCATGC TCATCTGGAT GGGCA 25

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAAAAGCTTA AAATGGCCTC GGTCGCC 27

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTTTCTAGAC TACATCGTTG TCTT 24

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AAAAAGCTTA AAATGAATCT GATCAAC 27

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAATCTAGAT TAGTCGAAAC TGAA 24

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AAAAAGCTTA AAATGCCTGG CGGA 24

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AAATCTAGAG GCTACAGGAA GTCC 24

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGGGGTACCA AAATGTCGGG GTGTGAT 27

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTTTTCTAGA TCAAGAGTTA TCATC 25

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1529 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Asx	Asp	Asp	Ala	His	Asx	Asp	Asp	Ala	Asx	His	Ala	Asx	His	Ala	Asp
1				5					10					15	
Asx	Ala	Asp	Asx	His	Asp	His	Ala	Ala	His	Ala	Asx	Asp	Ala	Asx	His
			20					25					30		
Asx	Ala	Asx	His	Ala	His	Ala	Asp	Asp	Asp	Ala	Ala	His	Ala	His	His
		35					40					45			
Ala	Ala	Asp	Asx	His	Asp	Asp	His	Ala	Asx	Asx	Asp	Ala	Asp	Asx	His
	50					55					60				
Asx	Asp	Asp	Ala	His	Asx	Asx	Ala	Asx	His	Ala	Asp	His	Ala	Ala	Asx
65					70					75					80
Asp	Asp	Asx	Asx	Asp	Asx	Asx	Ala	Asp	His	Asp	His	Asp	Asp	His	Asp
				85					90					95	
Asp	Ala	Ala	His	His	Asx	Asp	Asp	Ala	Ala	Ala	His	Asp	Asp	His	Ala
			100					105					110		
Ala	His	Ala	Ala	His	Asx	Ala	Ala	Asx	Asx	Asp	Ala	His	Asx	Asp	Ala
		115					120					125			
Ala	Asx	Ala	Asx	Asx	His	Ala	His	Asp	Asx	Asx	Asp	His	His	Asp	Ala
	130					135					140				
Asp	Asx	Ala	Asp	Asp	Ala	Ala	Asp	Asx	Ala	His	His	His	Asx	Asx	Ala
145					150					155					160
Ala	Asp	Ala	Asp	Ala	Asx	Ala	Ala	Asp	His	Ala	Asx	Ala	Ala	His	Ala
				165					170					175	
His	His	Asp	His	Asx	His	Ala	Asx	His	Asp	Asp	Asx	His	Asx	Asp	His
			180					185					190		
Asx	Ala	His	His	Asx	His	His	Asp	His	His	Asp	Asp	Ala	His	His	Asx
		195					200					205			
Asp	Asp	Ala	Asp	His	His	Asx	His	His	Asx	His	Asp	Asx	Asx	Ala	His
	210					215					220				
Asp	Asp	Ala	Ala	His	Ala	His	Asp	His	His	Asx	Ala	His	His	Ala	Asx
225					230					235					240
His	Ala	His	Asx	Asp	Asx	Asx	Asx	Asx	His	Asp	Ala	Asp	His	Ala	His
				245					250					255	
His	Ala	His	His	Asp	His	Asp	Ala	Ala	His	His	Ala	His	His	Asp	Asp
			260					265					270		
His	His	Asx	Ala	Ala	Ala	Asx	Asx	Asp	Asp	Ala	His	Asp	Asp	Asx	Asp
		275					280					285			
His	Asp	Asp	Ala	Asp	Ala	Asx	Ala	His	Asp	Asp	His	Ala	His	His	Asx
	290					295					300				
Asp	Ala	Ala	Ala	Asp	Ala	Ala	His	His	Asx	Ala	His	Asp	Asp	Asp	Ala
305					310					315					320
His	Asx	His	His	His	Asp	Ala	Asx	Asp	Ala	His	His	Asp	Asp	Asx	His
				325					330					335	
Asx	Ala	Asx	Ala	Ala	Asx	His	His	Asx	Asx	Ala	Ala	Ala	Asx	Asp	Asx
			340					345					350		

Ala Ala Asp Asx Ala His His Ala Ala His Asp His His His His Asx
355 360 365

Ala Ala Asx Asx His Asp His His Asx Asx His Asx Ala His His Ala
370 375 380

His His Asp Asx His Asp Asp His Asx Asx Asx Asx His Asp Ala His
385 390 395 400

Asx His Ala Asx Asx Asp Asx Asp His Asx His His His Asp Asx His
405 410 415

Asx Asx Asp Asp His His His Asp Asx His His Asx Ala Ala Asx Ala
420 425 430

His Asx Asp His Asx Ala Ala Asx Asx His Asp Ala Asx Ala Ala His
435 440 445

Asx Ala His His Asx His Asx Ala His Asx Asx His Asx Asp His Asx
450 455 460

Ala His His Asp His His Asx His Asp Asp Ala Asp Asx Asx Asx Ala
465 470 475 480

Asx His Asp Ala Ala Asp Ala His His Asx Asx Ala His Asp His Asx
485 490 495

Asx His Asp Asp His His His His His Asx His Asp Asp Asp His Ala
500 505 510

Ala Asx His Asx His His Asp Asp Ala Ala His Asp Asp Asx Asp Ala
515 520 525

Asx His His Asx Ala Ala His Asx Ala Ala His His His His Ala Asp
530 535 540

Asx Ala Ala His Asp Asp Asp Asx His Ala His Ala His Asp Ala Ala
545 550 555 560

Ala Ala Asx His Asx Asp Asp His His His Ala His Asp Asp Ala Asp
565 570 575

His His Asp Asp His Asp Asp Asx Asp Ala His His His His Asx Asx
580 585 590

Asp Asx Ala Asx Ala Asx Asx His Ala Asx Ala His His Asp Asp Asx
595 600 605

Asp Asx His Asx His Asx His His Asp Ala His His Asp Asp Ala Ala
610 615 620

Ala Asx Ala Ala Asx Ala His His His Asp Asx Asp Asp Ala His His
625 630 635 640

Asp Asx His Asp Ala His Ala Ala Asx Asp Asp His His Asp His Asp
645 650 655

Ala Ala Ala Ala His Asx Asp Asp Ala Asp His Asp Ala Asx Asx His
660 665 670

Ala His His His His Asx His Asp Ala Ala His Asp Ala His Asp Ala
675 680 685

Asp Asx Asx His Ala Ala Ala Asx His His Asp His His Asp Asx Ala
690 695 700

Ala His Asx Asp His Asx His Ala His His His Asx Asp Asp Asx Ala
705 710 715 720

His Ala His Asx Asp His His Asp Asp His Asp Ala His Asx Asx His
725 730 735

His Asx His Asp Asp His Asp His Asp His Asp Asx Ala Ala His His
740 745 750

Asp Asx Ala Asx His His His His Asx His His His Ala His Asx Ala
755 760 765

Asx Ala Ala Ala Asp Asx Ala Ala Asp Ala His His His Asx His Ala
770 775 780

Asx Asx Ala Asx His Ala His Asx Ala Asx Asx Ala His Asx Ala Ala
785 790 795 800

Ala Ala Ala Asp Asp Ala Ala His Asp Asp Ala Ala Ala His His Asx
805 810 815

Asp Asx Asp Ala Ala Ala Ala Asp Asp Asx Asp Asp Ala Ala Ala Asx
820 825 830

Asx Asp Ala Asx Ala Asp Ala Asx Asx Asp His Asx His Asx Asx Ala
835 840 845

His Asx Asx Ala His His Asx His His His Asp Asp Ala Asx Asx Ala
850 855 860

Asx Ala His His Asx Ala Asx Ala Ala Ala Asx His Asp His His Ala
865 870 875 880

His Asp Asp Asp Asx Ala Ala Asx His Asx His His Asx Ala Ala His
885 890 895

Asp His His His Asp Asp His His Asx His Asp Asx His His His Asp
900 905 910

Asx Asx Asp His His Ala Asx His Asx His Asx Ala Asx Ala Ala His
915 920 925

Asx His His Asx Asx Asx His Asp His His Ala His Asp Ala His Asp
930 935 940

Ala Asx Asx Asp His His Ala Asx Asx Ala Asx His Asx Asp His Asp
945 950 955 960

Asp Ala Asp Ala His His Asx Asx Asp Asp Asx His His Asx Asx His
965 970 975

Ala Ala Ala Asx Ala Ala Ala Ala His His Ala His Asp His Asx His
980 985 990

Asp Ala Ala Ala Ala Asx Asp Ala His Asp Ala Ala Ala His Asx His
995 1000 1005

Ala Asx Ala Asx His His His Asp Asx His Asx Ala Asx Ala Ala Asp
1010 1015 1020

His His His Asx Asx His Asx His His Asx His His Asx Ala Ala His
1025 1030 1035 1040

His His Asp His His Asx Asp Asx His Asp Asx Asp Ala His His Asp
1045 1050 1055

Asp Ala His Asx Asx Ala His Ala Asp His His Asp Asx His His Asx
1060 1065 1070

Asx Ala Ala Asp Ala His His Asx Ala Asx His Asp Asp Asx Asx Asp
1075 1080 1085

Ala Asx Ala Asx Asx Asx Asx Asp His His Ala Asx Asx His Asx Ala
1090 1095 1100

Ala Ala His His His Asp Asx Asx Ala His Ala Ala His Asx His His
1105 1110 1115 1120

Asp Asx Asp His Asp Asx His Asx His His His His Asx Ala His His
1125 1130 1135

Asx Asx Ala His His Asx His His Asx His His Asx His His Asx His
1140 1145 1150

Asp Asx Ala Ala Asx His Ala His Asx Asp His Asp His Asx Asx Ala
1155 1160 1165

Asp Ala Asx Asp Asx Asp His Asp Asx His His Ala His Asx Asx His
1170 1175 1180

Asp His His His His Asx His His His Asp Ala Asp His Asx His Ala
1185 1190 1195 1200

Asx His Asp Ala Asx Ala His His His His His Asp His Asp Ala His
1205 1210 1215

His Asp Asp His Asp Asp Ala Ala His His Asp Asx Asx Ala His Asp
1220 1225 1230

His Asx His His His His His Asx Ala Asx Ala His Asp Asp Ala His
1235 1240 1245

Ala Asx Asx His Asx Ala Asp Asx Asp Asx His Asx His Asp Asp Asx
1250 1255 1260

Ala Ala His Asp Asp Asp Ala His Ala Asx Ala Asx His Asx Asx Ala
1265 1270 1275 1280

Ala Ala Asx Asp His Asx Asp His Asp Asx Asx Ala His Asx His Asx
1285 1290 1295

Ala Asx His Ala Asx His Asx Ala Ala Asp Ala His His His Asp Asx
1300 1305 1310

Asx Asp Asx His Asx Ala Asp Asx His His His Asx Asx Asp His His
1315 1320 1325

His Asp Asx Ala Asx His Asx His His Ala His Asp Asp His His Asp
1330 1335 1340

Asp Asx Asx His His Asx His Asx Ala Asx Asx Asp Asp His Asp Asp
1345 1350 1355 1360

Asx Asx His Asp His Asp Asp Asx Asx Asx Asp His His Asp His His
1365 1370 1375

Ala His His Asp Ala Asp Asx Ala Asx His His Asx Asp His Asp Asp
1380 1385 1390

Ala Asx Ala Ala Asp Asx Asx Ala Ala Asp His Ala His Asx His His
1395 1400 1405

Ala His Ala Ala Ala His Ala His His His Ala His Ala Asp Asx Ala
1410 1415 1420

His His Ala Asp Ala Asp His Ala His Ala Asx His His Asp His His
1425 1430 1435 1440

Ala His Ala His Asp His His Asp His His His His His Ala His His
1445 1450 1455

Ala Ala Asp Asx His Asp His Asp Asp Ala Ala His Ala Ala Ala Ala
1460 1465 1470

His Ala Ala His His Ala His His Ala Ala Ala Ala Ala Ala Ala Ala
1475 1480 1485

Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Asp Asx Asp
1490 1495 1500

Asx Asx Asp Asx His Asx Asp Ala Asp Asx Ala His His Asx Ala His
1505 1510 1515 1520

Asp Asp Ala Asp Ala Asp Ala Ala Ala
1525

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Ser Pro Asn Arg Trp Ile Leu Leu Leu Ile Phe Tyr Ile Ser Tyr
1 5 10 15

Leu Met Phe Gly Ala Ala Ile Tyr Tyr His Ile Glu His Gly Glu Glu
20 25 30

Lys Ile Ser Arg Ala Glu Gln Arg Lys Ala Gln Ile Ala Ile Asn Glu
35 40 45

Tyr Leu Leu Glu Glu Leu Gly Asp Lys Asn Thr Thr Thr Gln Asp Glu
50 55 60

Ile Leu Gln Arg Ile Ser Asp Tyr Cys Asp Lys Pro Val Thr Leu Pro
65 70 75 80

Pro Thr Tyr Asp Asp Thr Pro Tyr Thr Trp Thr Phe Tyr His Ala Phe
85 90 95

Phe Phe Ala Phe Thr Val Cys Ser Thr Val Gly Tyr Gly Asn Ile Ser
100 105 110

Pro Thr Thr Phe Ala Gly Arg Met Ile Met Ile Ala Tyr Ser Val Ile
115 120 125

Gly Ile Pro Val Asn Gly Ile Leu Phe Ala Gly Leu Gly Glu Tyr Phe
130 135 140

Gly Arg Thr Phe Glu Ala Ile Tyr Arg Arg Tyr Lys Lys Tyr Lys Met
145 150 155 160

Ser Thr Asp Met His Tyr Val Pro Pro Gln Leu Gly Leu Ile Thr Thr
165 170 175

Val Val Ile Ala Leu Ile Pro Gly Ile Ala Leu Phe Leu Val Leu Pro
180 185 190

Cys Val Gly Val His Leu Leu Arg Glu Leu Gly Leu Ser Ser Ile Ser
195 200 205

L u Tyr Tyr Ser Tyr Val Thr Ile Thr Thr Ile Gly Phe Gly Asp Tyr
210 215 220

Val Pro Thr Phe Gly Ala Asn Gln Pro Lys Glu Phe Gly Gly Trp Phe
225 230 235 240

Val Val Tyr Gln Ile Ph Val Ile Val Trp Phe Ile Phe Ser Leu Gly
245 250 255

Tyr Leu Val Met Ile Met Thr Phe Ile Thr Arg Gly Leu Gln Ser Lys
260 265 270

Lys Leu Ala Tyr Leu Glu Gln Gln Leu Ser Ser Asn Leu Lys Ala Thr
275 280 285

Gln Asn Arg Ile Trp Ser Gly Val Thr Lys Asp Val Gly Tyr Leu Arg
290 295 300

Arg Met Leu Asn Glu Leu Tyr Ile Leu Lys Val Lys Pro Val Tyr Thr
305 310 315 320

Asp Val Asp Ile Ala Tyr Thr Leu Pro Arg Ser Asn Ser Pro Leu Ser
325 330 335

Met Tyr Arg Val Glu Pro Ala Pro Ile Pro Ser Arg Lys Arg Ala Phe
340 345 350

Ser Val Cys Ala Asp Met Val Gly Ala Gln Arg Glu Ala Gly Met Val
355 360 365

His Ala Asn Ser Asp Thr Asp Leu Thr Lys Leu Asp Arg Glu Lys Thr
370 375 380

Phe Glu Thr Ala Glu Ala Tyr His Gln Thr Thr Asp Leu Leu Ala Lys
385 390 395 400

Val Val Asn Ala Leu Ala Thr Val Lys Pro Pro Pro Ala Leu Gln Glu
405 410 415

Asp Ala Ala Leu Tyr Gly Gly Tyr His Gly Phe Ser Asp Ser Gln Ile
420 425 430

Leu Ala Ser Glu Trp Ser Phe Ser Thr Val Asn Glu Phe Thr Ser Pro
435 440 445

Arg Arg Pro Arg Ala Arg Ala Cys Ser Asp Phe Asn Leu Glu Ala Pro
450 455 460

Arg Trp Gln Ser Glu Arg Pro Leu Arg Ser Ser His Asn Glu Trp
465 470 475

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 335 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Ser Asp Gln Leu Phe Val Ala Phe Glu Lys Tyr Phe Leu Thr Ser
1 5 10 15

Asn Glu Val Lys Lys Asn Ala Ala Thr Glu Thr Trp Thr Phe Ser Ser
20 25 30

Ser Ile Phe Phe Ala Val Thr Val Val Thr Thr Ile Gly Tyr Gly Asn
35 40 45

Pro Val Pro Val Thr Asn Ile Gly Arg Ile Trp Ile Leu Phe Ser Leu
50 55 60

Ile Gly Ile Pro Leu Thr Leu Val Thr Ile Ala Leu Ala Gly Lys Phe
65 70 75 80

Leu Ser Glu His Leu Val Trp Leu Tyr Gly Asn Tyr Leu Lys Leu Lys
85 90 95

Tyr Leu Ile Leu Ser Arg His Arg Lys Glu Arg Arg Glu His Val Cys
100 105 110

Glu His Cys His Ser His Gly Met Gly His Asp Met Asn Ile Glu Glu
115 120 125

Lys Arg Ile Pro Ala Phe Leu Val Leu Ala Ile Leu Ile Val Tyr Thr
130 135 140

Ala Phe Gly Gly Val Leu Met Ser Lys Leu Glu Pro Trp Ser Phe Phe
145 150 155 160

Thr Ser Phe Tyr Trp Ser Phe Ile Thr Met Thr Thr Val Gly Phe Gly
165 170 175

Asp Leu Met Pro Arg Arg Asp Gly Tyr Met Tyr Ile Ile Leu Leu Tyr
180 185 190

Ile Ile Leu Gly Lys Phe Ser Met Lys Lys Lys Gln Lys Phe Lys Ile
195 200 205

Phe Leu Gly Leu Ala Ile Thr Thr Met Cys Ile Asp Leu Val Gly Val
210 215 220

Gln Tyr Ile Arg Lys Ile His Tyr Phe Gly Arg Lys Ile Gln Asp Ala
225 230 235 240

Arg Ser Ala Leu Ala Val Val Gly Gly Lys Val Val Leu Val Ser Glu
245 250 255

Leu Tyr Ala Asn Leu Met Gln Lys Arg Ala Arg Asn Met Ser Arg Glu
260 265 270

Ala Phe Ile Val Glu Asn Leu Tyr Val Ser Lys His Ile Ile Pro Phe
275 280 285

Ile Pro Thr Asp Ile Arg Cys Ile Arg Tyr Ile Asp Gln Thr Ala Asp
290 295 300

Ala Ala Thr Ile Ser Thr Ser Ser Ser Ala Ile Asp Met Gln Ser Cys
305 310 315 320

Arg Phe Cys His Ser Arg Tyr Ser Leu Asn Arg Ala Phe Lys Xaa
325 330 335

What is claimed is:

1. A potassium channel comprising four hydrophobic domains capable of forming transmembrane helices, wherein
   (i) a first pore-forming domain is interposed between a first and a second transmembrane helix; and
   (ii) a second pore-forming domain is interposed between a third and a fourth transmembrane helix.

2. The potassium channel of claim 1 wherein each pore-forming domain comprises a potassium selective peptide motif.

3. The potassium channel of claim 2 wherein the peptide motif is selected from the group consisting of a Y/F-G dipeptide motif and a G-Y/F-G tripeptide motif.

4. The potassium channel of claim 3 wherein at least one pore-forming domain is positioned proximal to an exterior portion of a cell membrane.

5. The potassium channel of claim 4 further comprising an amino-terminal glycosylation site.

6. The potassium channel of claim 5 wherein said glycosylation site is asparagine-linked.

7. The potassium channel of claim 6 characterized in that it belongs to a class of invertebrates.

8. The potassium channel of claim 7 characterized in that

it is insect-derived.

9. The potassium channel of claim 7 characterized in that it is nematode-derived.

10. An isolated nucleotide sequence capable of encoding DmORF-1.

11. The isolated nucleotide sequence of Claim 10 comprising the nucleotide sequence depicted in Seq. I.D. No. 1.

12. An isolated nucleotide sequence capable of encoding CORK.

13. The isolated nucleotide sequence of Claim 12 encoding for the protein depicted in Sequence I.D. No. 36.

14. An expression vector capable of expressing a heterologous potassium channel in a cell membrane of a yeast cell comprising the nucleotide sequence of Claim 10.

15. An expression vector capable of expressing a heterologous, potassium channel in a cell membrane of a yeast cell comprising the nucleotide sequence of Claim 11.

16. An expression vector capable of expressing a heterologous potassium channel in a cell membrane of a yeast cell comprising the nucleotide sequence of Claim 12.

17. An expression vector capable of expressing a heterologous potassium channel in a cell membrane of a yeast cell wherein the potassium channel comprises the amino acid sequence of Claim 13.

18. A transformed yeast cell comprising the nucleotide sequences of Claims 10, 11, 12 or 13.

19. A transformed yeast cell comprising the expression vector of claims 14, 15, 16 or 17.

20. A method of assaying substances to determine effects on cell growth, the method comprising the steps of:

a. preparing cultures of yeast cells in liquid medium lacking uracil, the liquid medium consisting of a concentration of KCl adequate to support growth of potassium-dependent mutant strains;

b. plating the yeast cells in uracil-free agar medium, the agar medium consisting of sufficient KCl to selectively support growth of potassium-dependent mutant strains containing a heterologous potassium channel of claim 1;

c. applying substances to the agar plate;

d. incubating the agar plate to permit growth; and

e. identifying zones of growth around the substances, wherein the level of growth indicates whether or not activity of the heterologous potassium channel has been modulated as compared to control.

21. The yeast cell of Claim 20 further comprising a nucleotide sequence encoding RAK, or a nucleotide sequence of Claim 10, 11, 12 or 13.

22. The method of claim 20, wherein said effect on cell

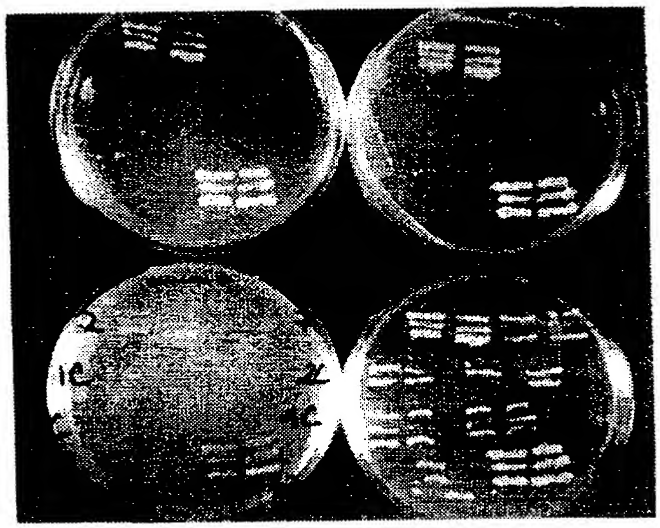
growth is modulated by activation of the potassium channel.

23. The method of claim 20, wherein said effect on cell growth is modulated by inhibition of said potassium channel.

24. A method of selectively inhibiting insect pests by applying to such insect pests a substance capable of inhibiting a potassium channel substantially homologous to that encoded by the nucleotide sequence of claim 10.

25. A method of selectively inhibiting nematode pests by applying to such pests a substance capable of inhibiting a potassium channel substantially homologous to that encoded by the nucleotide sequence of claim 12.

26. A method of modulating the activity of a potassium channel positioned in a cellular membrane and comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, and SEQ ID NO: 36, by contacting said cellular membrane with a substance, in an amount and for a period of time sufficient to inhibit the ability of potassium ions to pass through said channel.



SC galactose, 100 mM KCl

SC glucose, 0mM KCl

SC galactose, 0 mM KCl

SC glucose, 100 mM KCl

FIG. 1

```
5'...ACGCGATCGCCGCGAGTGTATATTTTTTTTTAGCTCAGTCTTCAGTGTTTCGCGATTCTCTTTAAAGAAAAAAAAAATAATAAGTCAA
AACTACAAACCACACAGCGAAAGGCGAAAGCAACGGTTCCTGCGAGTGTTTATTTTTTTTTCAACAATTTTTGATCGTAGTGCGACAATCCGTCGAGC   -1
                                                                 M1               20
                                  10
Met Ser Pro Asn Arg Trp Ile Leu Leu Leu Ile Phe Tyr Ile Ser Tyr Leu Met Phe Gly Ala Ala Ile Tyr Tyr
ATG TCG CCG AAT CGA TGG ATC CTG CTG CTC ATC TTC TAC ATA TCC TAC CTG ATG TTC GGG GCG GCA ATC TAT TAC   75
                                                        40                                      50
                    30
His Ile Glu His Gly Glu Glu Lys Ile Ser Arg Ala Glu Gln Arg Lys Ala Gln Ile Ala Ile Asn Glu Tyr Leu
CAT ATT GAG CAC GGC GAG GAG AAG ATA TCG CGC GCC GAA CAG CGC AAG GCG CAA ATT GCA ATC AAC GAA TAT CTG   150
                                                                          70
                                  G       60
Leu Glu Glu Leu Gly Asp Lys Asn Thr Thr Thr Gln Asp Glu Ile Leu Gln Arg Ile Ser Asp Tyr Cys Asp Lys
CTG GAG GAG CTG GGC GAC AAG AAT ACG ACC ACA CAG GAT GAG ATT CTT CAA CGG ATC TCG GAT TAC TGT GAC AAA   225
                                                       90                                      100
                    80
Pro Val Thr Leu Pro Pro Thr Tyr Asp Asp Thr Pro Tyr Thr Trp Thr Phe Tyr His Ala Phe Phe Phe Ala Phe
CCG GTT ACA TTG CCG CCG ACA TAT GAT GAT ACG CCC TAC ACG TGG ACC TTC TAC CAT GCC TTC TTC TTC GCC TTC   300
                                                                              120
         H5-1                            110
Thr Val Cys Ser Thr Val Gly Tyr Gly Asn Ile Ser Pro Thr Thr Phe Ala Gly Arg Met Ile Met Ile Ala Tyr
ACC GTT TGC TCC ACG GTG GGA TAT GGG AAT ATA TCG CCA ACC ACC TTC GCC GGA CGG ATG ATC ATG ATC GCG TAT   375
                                                        140                                     150
                130    M2
Ser Val Ile Gly Ile Pro Val Asn Gly Ile Leu Phe Ala Gly Leu Gly Glu Tyr Phe Gly Arg Thr Phe Glu Ala
TCG GTG ATT GGC ATC CCC GTC AAT GGT ATC CTC TTT GCC GGC CTC GGC GAA TAC TTT GGA CGT ACG TTT GAA GCG   450
                                                                          170
                                         160
Ile Tyr Arg Arg Tyr Lys Lys Tyr Lys Met Ser Thr Asp Met His Tyr Val Pro Pro Gln Leu Gly Leu Ile Thr
ATC TAC AGA CGC TAC AAA AAG TAC AAG ATG TCC ACG GAT ATG CAC TAT GTC CCG CCG CAG CTG GGA TTG ATC ACC   525
                                                        190                                     200
                180    M3
Thr Val Val Ile Ala Leu Ile Pro Gly Ile Ala Leu Phe Leu Val Leu Pro Cys Val Gly Val His Leu Leu Arg
ACG GTG GTG ATT GCC CTG ATT CCG GGA ATA GCT CTC TTC CTG GTG CTG CCC TGC GTG GGT GTT CAC CTA CTT CGA   600
                                                            H5-2              220
                                         210
Glu Leu Gly Leu Ser Ser Ile Ser Leu Tyr Tyr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe Gly Asp Tyr Val
GAA CTG GGC CTA TCT TCC ATC TCG CTG TAC TAC AGC TAT GTG ACC ACC ACA ACA ATT GGA TTC GGT GAC TAT GTG   675
                                                        240                                     250
                230
Pro Thr Phe Gly Ala Asn Gln Pro Lys Glu Phe Gly Gly Trp Phe Val Val Tyr Gln Ile Phe Val Ile Val Trp
CCC ACA TTT GGA GCC AAC CAG CCC AAG GAG TTC GGC GGC TGG TTC GTG GTC TAT CAG ATC TTT GTG ATC GTG TGG   750
                                         260                                  270
    M4
Phe Ile Phe Ser Leu Gly Tyr Leu Val Met Ile Met Thr Phe Ile Thr Arg Gly Leu Gln Ser Lys Lys Leu Ala
TTC ATC TTC TCG CTG GGA TAT CTT GTG ATG ATC ATG ACA TTT ATC ACT CGG GGC CTC CAG AGC AAG AAG CTG GCA   825
                                                        290                                     300
                280
Tyr Leu Glu Gln Gln Leu Ser Ser Asn Leu Lys Ala Thr Gln Asn Arg Ile Trp Ser Gly Val Thr Lys Asp Val
TAC CTG GAG CAG CAG TTG TCC TCC AAC CTG AAG GCC ACA CAG AAT CGC ATC TGG TCT GGC GTC ACC AAG GAT GTG   900
                                                                          320
                                         310
Gly Tyr Leu Arg Arg Met Leu Asn Glu Leu Tyr Ile Leu Lys Val Lys Pro Val Tyr Thr Asp Val Asp Ile Ala
GGC TAC CTC CGG CGA ATG CTC AAC GAG CTG TAC ATC CTC AAA GTG AAG CCT GTG TAC ACC GAT GTA GAT ATC GCC   975
```

FIG. 2A

```
                  330                                          340                                          350
Tyr Thr Leu Pro Arg Ser Asn Ser Cys Pro Asp Leu Ser Met Tyr Arg Val Glu Pro Ala Pro Ile Pro Ser Arg
TAC ACA CTG CCA CGT TCC AAT TCG TGT CCG GAT CTG AGC ATG TAC CGC GTG GAG CCG GCT CCC ATT CCC AGC CGG   1050
                                    360                                         370
Lys Arg Ala Phe Ser Val Cys Ala Asp Met Val Gly Ala Gln Arg Glu Ala Gly Met Val His Ala Asn Ser Asp
AAG AGG GCA TTC TCC GTG TGC GCC GAC ATG GTT GGC GCC CAA AGG GAG GCG GGC ATG GTA CAC GCC AAT TCC GAT   1125
                380                                         390                                     400
Thr Asp Leu Thr Lys Leu Asp Arg Glu Lys Thr Phe Glu Thr Ala Glu Ala Tyr His Gln Thr Thr Asp Leu Leu
ACG GAT CTA ACC AAA CTG GAT CGC GAG AAG ACA TTC GAG ACG GCG GAG GCG TAC CAC CAG ACC ACC GAT TTG CTG   1200
                                    410                                         420
Ala Lys Val Val Asn Ala Leu Ala Thr Val Lys Pro Pro Pro Ala Glu Gln Glu Asp Ala Ala Leu Tyr Gly Gly
GCC AAG GTG GTC AAC GCA CTG GCC ACG GTG AAG CCA CCG CCG GCG GAA CAG GAA GAT GCG GCT CTC TAT GGT GGC   1275
                430                                         440                                     450
Tyr His Gly Phe Ser Asp Ser Gln Ile Leu Ala Ser Glu Trp Ser Phe Ser Thr Val Asn Glu Phe Thr Ser Pro
TAT CAT GGC TTC TCC GAC TCC CAG ATC CTG GCC AGC GAA TGG TCG TTC TCG ACG GTC AAC GAG TTC ACA TCA CCG   1350
                                    460                                         470
Arg Arg Pro Arg Ala Arg Ala Cys Ser Asp Phe Asn Leu Glu Ala Pro Arg Trp Gln Ser Glu Arg Pro Leu Arg
CGA CGT CCA AGA GCA CGT GCC TGC TCC GAT TTC AAT CTG GAG GCA CCT CGC TGG CAG AGC GAG AGG CCA CTG CGT   1425
                480                                         490                                     500
Ser Ser His Asn Glu Trp Thr Trp Ser Gly Asp Asn Gln Gln Ile Gln Glu Ala Phe Asn Gln Arg Tyr Lys Gly
TCG AGC CAC AAC GAA TGG ACA TGG AGC GGC GAC AAC CAG CAG ATC CAG GAG GCA TTC AAC CAG CGC TAC AAG GGA   1500
                                    510                                         520
Gln Gln Arg Ala Asn Gly Ala Ala Asn Ser Thr Met Val His Leu Glu Pro Asp Ala Leu Glu Glu Gln Leu Arg
CAG CAG CGT GCC AAC GGA GCA GCC AAC TCG ACC ATG GTC CAT CTG GAG CCG GAT GCT TTG GAG GAG CAG CTG AGA   1575
                530                                         540                                     550
Asn Asn His Arg Val Pro Val Ala Ser Arg Ser Ser Pro Cys Arg Met Val Cys Asp Val Cys Phe Pro Ser Arg
AAC AAT CAC CGG GTG CCG GTC GCG TCA AGA AGT TCT CCA TGC CGG ATG GTC TGC GAC GTC TGT TTC CCT TCC AGA   1650
                                    560                                         570
Arg Ser Thr Pro Arg Arg Ile Trp Ser Ala Ser Cys Pro Trp Ser Arg Tyr Pro Arg Val Ser Ser Arg Arg Lys
AGA AGC ACC CCT CGC AGG ATC TGG AGC GCA AGT TGT CCG TGG TCT CGG TAC CCG AGG GTG TCA TCT CGC AGG AAG   1725
                580                                         590                                     600
Pro Asp Pro Arg Trp Thr Thr Thr Ser Thr Arg Ser Arg Arg Pro Pro Val Asn Pro Ile Cys Ala Thr Asp Ala
CCA GAT CCC CGC TGG ACT ACT ACA TCA ACA CGG TCA CGG CGG CCT CCA GTC AAT CCT ATT TGC GCA ACG GAC GCG   1800
                                    610                             618
Val Arg His Arg Pro Ser Asn Arg Met Ala Ala Trp Pro Ala Ala Ala Ala Gly
GTC CGC CAC CGC CCT TCG AAT CGA ATG GCA GCT TGG CCA GCG GCG GCG GCG GGC TAA CGAACATGGGCTTCCAGATGGAG   1880

GATGGAGCAACCCCGCCATCGGCATTGGGCGGTGGAGCCTATCAACGCAAGGCGGCTGCTGGCAAGCGCCGACGCGAGAGCATCTACACCCAGAATCAA
GCCCCATCCGCTCGCCGGGGCAGCATGTATCCGCCGACCGCGCACGCCTTGGCCCAGATGCAGATGCGACGCGGCAGCTTGGCAACCAGTGGCTCTGGA
TCGGCGGCCATGGCGGCAGTGGCCGCGCGTCGTGGCAGCCTCTTCCCAGCTACAGCATCGGCATCATCGCTGACCTCTGCTCCGCGCCGAAGCAGCATA
TTCTCGGTTACCTCCGAAAAGGATATGAATGTGCTGGAGCAGACGACCATTGCGGATCTGATTCGTGCGCTCGAG . . .3'
```

FIG. 2B

```
                                        10                                      20
Met Ser Asp Gln Leu Phe Val Ala Phe Glu Lys Tyr Phe Leu Thr Ser Asn Glu Val Lys
ATG TCC GAT CAG CTG TTT GTC GCA TTT GAG AAG TAT TTC TTG ACG AGT AAC GAG GTC AAG   60
                                        30                                      40
Lys Asn Ala Ala thr Glu Thr Trp Thr Phe Ser Ser Ser Ile Phe Phe Ala Val Thr Val
AAG AAT GCA GCA ACG GAG ACA TGG ACA TTT TCA TCG TCC ATT TTC TTT GCC GTA ACC GTC  120
   H5-1                                 50                                      60
Val Thr Thr Ile Gly Tyr Gly Asn Pro Val Pro Val Thr Asn Ile Gly Arg Ile Trp Cys
GTC ACT ACC ATC GGA TAC GGT AAT CCA GTT CCA GTG ACA AAC ATT GGA CGG ATA TGG TGT  180
                        M2              70                                      80
Ile Leu Phe Ser Leu Leu Gly Ile Pro Leu Thr Leu Val Thr Ile Ala Asp Leu Ala Gly
ATA TTG TTC TCC TTG CTT GGA ATA CCT CTA ACA CTG GTT ACC ATC GCT GAC TTG GCA GGT  240
                                        90                                     100
Lys Phe Leu Ser Glu His Leu Val Trp Leu Tyr Gly Asn Tyr Leu Lys Leu Lys Tyr Leu
AAA TTC CTA TCT GAA CAT CTT GTT TGG TTG TAT GGA AAC TAT TTG AAA TTA AAA TAT CTC  300
                                       110                                     120
Ile Leu Ser Arg His Arg Lys Glu Arg Arg Glu His Val Cys Glu His Cys His Ser His
ATA TTG TCA CGA CAT CGA AAA GAA CGG AGA GAG CAC GTT TGT GAG CAC TGT CAC AGT CAT  360
                                       130                                     140
Gly Met Gly His Asp Met Asn Ile Glu Glu Lys Arg Ile Pro Ala Phe Leu Val Leu Ala
GGA ATG GGG CAT GAT ATG AAT ATC GAG GAG AAA AGA ATT CCT GCA TTC CTG GTA TTA GCT  420
            M3                         150                                     160
Ile Leu Ile Val Tyr Thr Ala Phe Gly Gly Val Leu Met Ser Lys Leu Glu Pro Trp Ser
ATT CTG ATA GTA TAT ACA GCG TTT GGC GGT GTC CTA ATG TCA AAA TTA GAG CCG TGG TCT  480
```

FIG. 3A

```
                                        170     H5-2                         180
Phe Phe Thr Ser Phe Tyr Trp Ser Phe Ile Thr Met Thr Thr Val Gly Phe Gly Asp Leu
TTC TTC ACT TCA TTC TAC TGG TCC TTC ATT ACA ATG ACT ACT GTC GGG TTT GGC GAC TTG    540
                                    190                                     200
Met Pro Arg Arg Asp Gly Tyr Met Tyr Ile Ile Leu Leu Tyr Ile Ile Leu Gly Lys Phe
ATG CCC AGA AGG GAC GGA TAC ATG TAT ATC ATA TTG CTC TAT ATC ATT TTA GGT AAA TTT    600
                                    210                             M4      220
Ser Met Lys Lys Lys Gln Lys Phe Lys Ile Phe Leu Gly Leu Ala Ile Thr Thr Met Cys
TCA ATG AAA AAA AAA CAA AAA TTC AAA ATA TTT TTA GGT CTT GCA ATA ACT ACA ATG TGC    660
                                    230                                     240
Ile Asp Leu Val Gly Val Gln Tyr Ile Arg Lys Ile His Tyr Phe Gly Arg Lys Ile Gln
ATT GAT TTG GTA GGA GTA CAG TAT ATT CGA AAG ATT CAT TAT TTC GGA AGA AAA ATT CAA    720
                                    250                                     260
Asp Ala Arg Ser Ala Leu Ala Val Val Gly Gly Lys Val Val Leu Val Ser Glu Leu Tyr
GAC GCT AGA TCT GCA TTG GCG GTT GTA GGA GGA AAG GTA GTC CTT GTA TCA GAA CTC TAC    780
                                    270                                     280
Ala Asn Leu Met Gln Lys Arg Ala Arg Asn Met Ser Arg Glu Ala Phe Ile Val Glu Asn
GCA AAT TTA ATG CAA AAG CGA GCT CGT AAC ATG TCC CGA GAA GCT TTT ATA GTG GAG AAT    840
                                    290                                     300
Leu Tyr Val Ser Lys His Ile Ile Pro Phe Ile Pro Thr Asp Ile Arg Cys Ile Arg Tyr
CTC TAT GTT TCC AAA CAC ATC ATA CCA TTC ATA CCA ACT GAT ATC CGA TGT ATT CGA TAT    900
                                    310                                     320
Ile Asp Gln Thr Ala Asp Ala Ala Thr Ile Ser Thr Ser Ser Ser Ala Ile Asp Met Gln
ATT GAT CAA ACT GCC GAT GCT GCT ACC ATT TCC ACG TCA TCG TCT GCA ATT GAT ATG CAA    960
                                    330                 336
Ser Cys Arg Phe Cys His Ser Arg Tyr Ser Leu Asn Arg Ala Phe Lys
AGT TGT AGA TTT TGT CAT TCA AGA TAT TCT CTC AAT CGT GCA TTC AAA TAG               1011
```

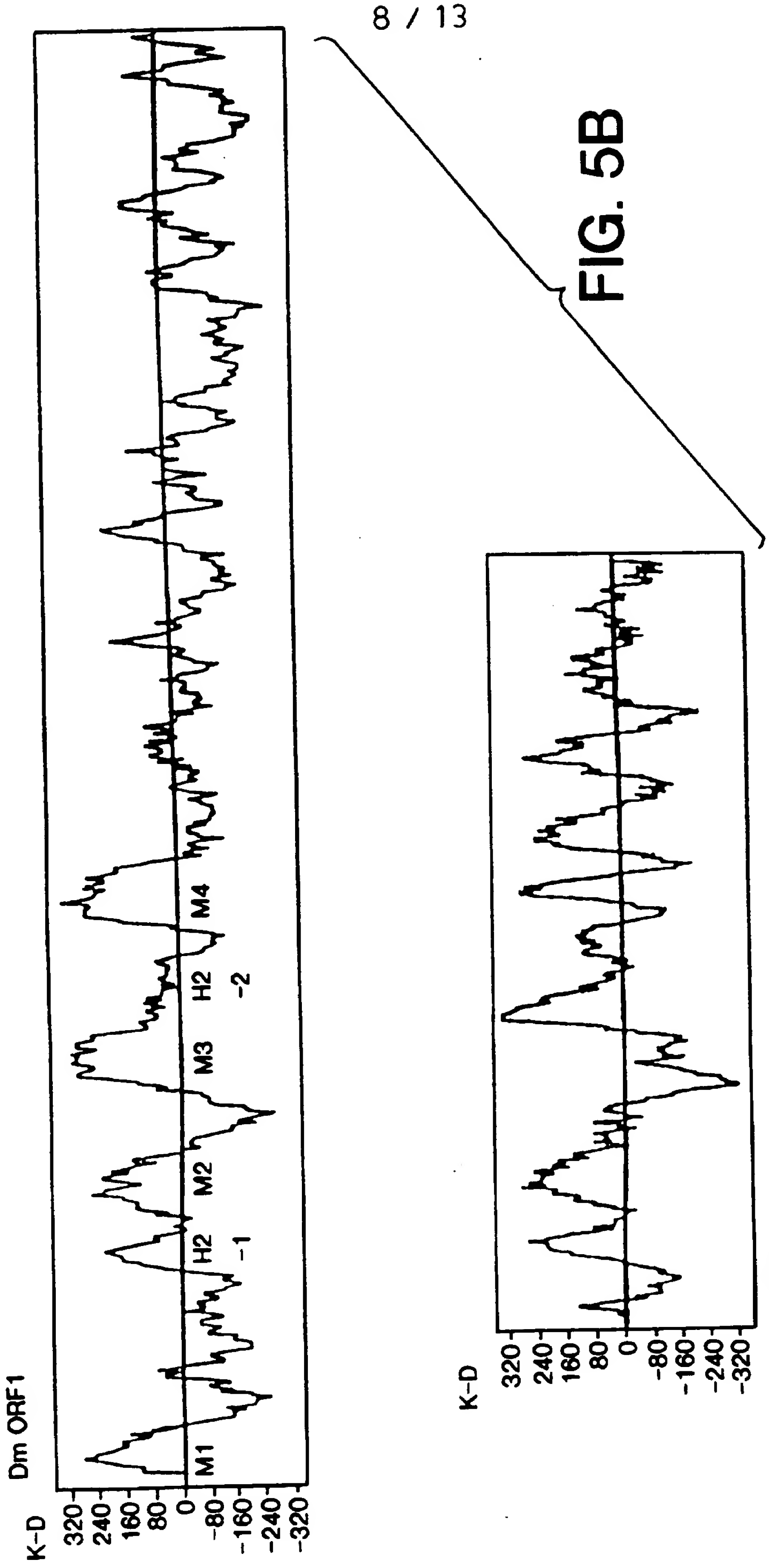
FIG. 3B

```
Ce orf1     ---------- ---------- ---------- ---------- ----------
Dm orf1     MSPNRWILLL IFYISYLMFG AAIYYHIEHG EEKISRAEQR KAQIAINEYL   50
Consensus   .......... .......... .......... .......... ..........   50

Ce orf1     ---------- --MSDQLFVA FEKYFLTSNE VKKNAATETW TFSSSIFFAV   38
Dm orf1     LEELGDKNTT TQDEILQRIS DYCDKPVTLP PTYDDTPYTW TFYHAFFFAF  100
Consensus   .......... .......... .......... ........TW TF...FFA.  100

Ce orf1     TVVTTIGYGN PVPVTNIGRI WCILFSLLGI PLTLVTIAAL AGKFLSEHLV   88
Dm orf1     TVCSTMGYGN ISPITFAGRM IMIAYSVIGI PVNGILFAAL ----------  140
Consensus   TV..T.GYGN ..P.T..GR. ..I..S..GI P......A.L ..........  150

Ce orf1     WLYGNYLKLK YLILSRHRKE RREHVCEHCH SHGMGHDMNI EEKRIPAFLV  138
Dm orf1     ---GEYFGRT FEAIYRRYKK YKMSTDMHYV PPQLGLITTV VIALIPGIAL  187
Consensus   ...G.Y.... .....R..K. .......H.. ....G..... ....IP....  200

Ce orf1     LAILIVYTAF GGVLMSKLEP WSFFTSFYWS FITYTTVGFG DLMPRRDGYM  188
Dm orf1     FLVLPCVGVH LLRELGLSS- ----ISLYMS YVTITTIGFG DYVPT-FGAN  231
Consensus   ...L...... .......... .....S.Y.S ..T.TT.GFG D..P...G..  250

Ce orf1     YIILLYIILG KFSMKKKQKF KIFLGLAITT MCIDLVGMQY IRKIHYFGRK  238
Dm orf1     QPKEFGGWFV VYQIFVIVWF IFSLGYLVMI MTFITRGLQS KKLAYLEQQL  281
Consensus   .......... .........F ...LG..... M.....G.Q. ..........  300

Ce orf1     IQDARSALAV VGGKVVLVSE LYANLMQKRA RNMSREAFIV ENLYVSKHII  288
Dm orf1     SSNLKATQNR IWSGVTKDVG YLRRMLNELY ILKVKPVYTD VDIAYTLPRS  331
Consensus   .......... ....V..... .......... .......... ..........  350

Ce orf1     PFIPTDIRCI -RYIDQTADA ATISTSSSAI DMQSCRFCHS RYSLNRAFKK  337
Dm orf1     NSCPDLSMYR VEPAPIPSRK RAFSVCADMV GAQREAGMVH ANSDTDLTKL  381
Consensus   ...P...... .......... ...S...... ..Q....... ..S.....K.  400

Ce orf1     ---------- ---------- ---------- ---------- ----------  337
Dm orf1     DREKTFETAE AYHQTTDLLA KVVNALATVK PPPAEQEDAA LYGGYHGFSD  431
Consensus   .......... .......... .......... .......... ..........  450

Ce orf1     ---------- ---------- ---------- ---------- ----------  337
Dm orf1     SQILASEWSF STVNEFTSPR RPRARACSDF NLEAPRWQSE RPLRSSHNEW  481
Consensus   .......... .......... .......... .......... ..........  500
```

FIG. 4

```
mIRK        AFLFSIETQTTIGYGFRCVTDECP     {G,A,S,T},   {D,E}
hROMK1      AFLFSLETQVTIGYGFRCVTEQCA     {N,Q} ,   {K,R,H}
rGIRK1      AFLFFIETEATIGYGYRYITDHCP     {F,Y,W}={I,L,M,V}
            ||.|..  .|.|||    |   .
Dm H5-1     AFFFAFTVCSTVGYGNISPTTFAG
            ||.||.|   ||||||||.||||||
Shak        AFWWAVVTMTTVGYGDMTPVGFWG
Shal        AFWYTIVTMTTLGYGDMVPETIAG
Shab        AFWWAGITMTTVGYGDICPTTALG
Shaw        GLWWALVTMTTVGYGDMAPKTYIG
Eag         ALYFTMTCMTSVGFGNVAAETDNE
Slo         CVYFLIVTMSTVGYGDVYCETVLG
            .||  |.||  |||||||.|||  .|
Dm H5-2     SLYTSYVTTTTTIGFGDYVPTFGAN

Dm H5-1     AFFFAFTVCSTVGYGNISPTTFAG
Ce 5-1      SIFFAVTVVTTIGYGNPVPVTNTG
Dm H5-2     SLYTSYVTTTTTIGFGDYVPTFGAN
Ce H5-2     SFYWSFITMTTVGFGDLMPRRDGY
```

FIG. 5A



Dm ORF1
K–D
320
240
160
80
0
–80
–160
–240
–320
M1
H2
–1
M2
M3
H2
–2
M4
Ce ORF1

FIG. 5B

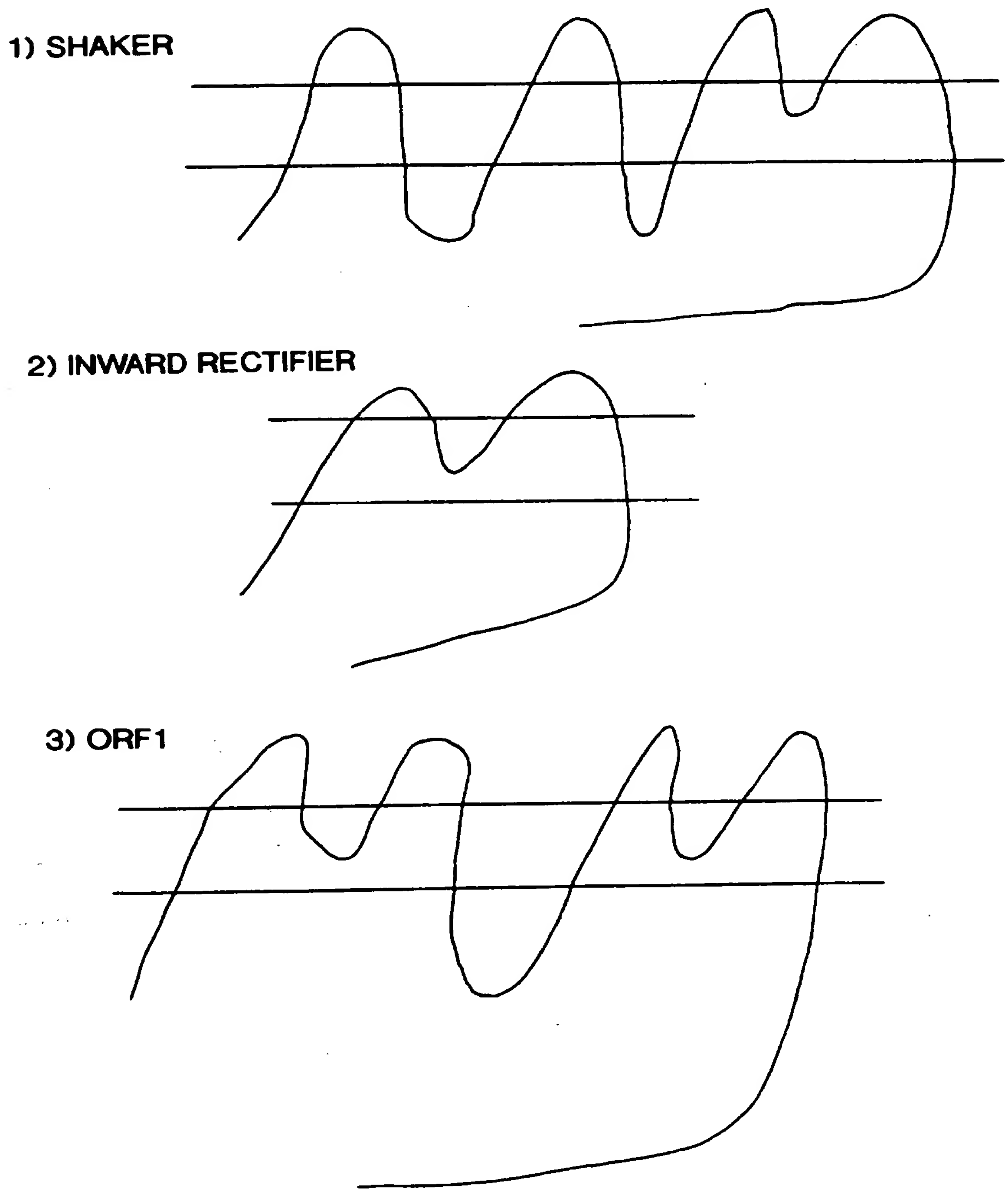


FIG. 6

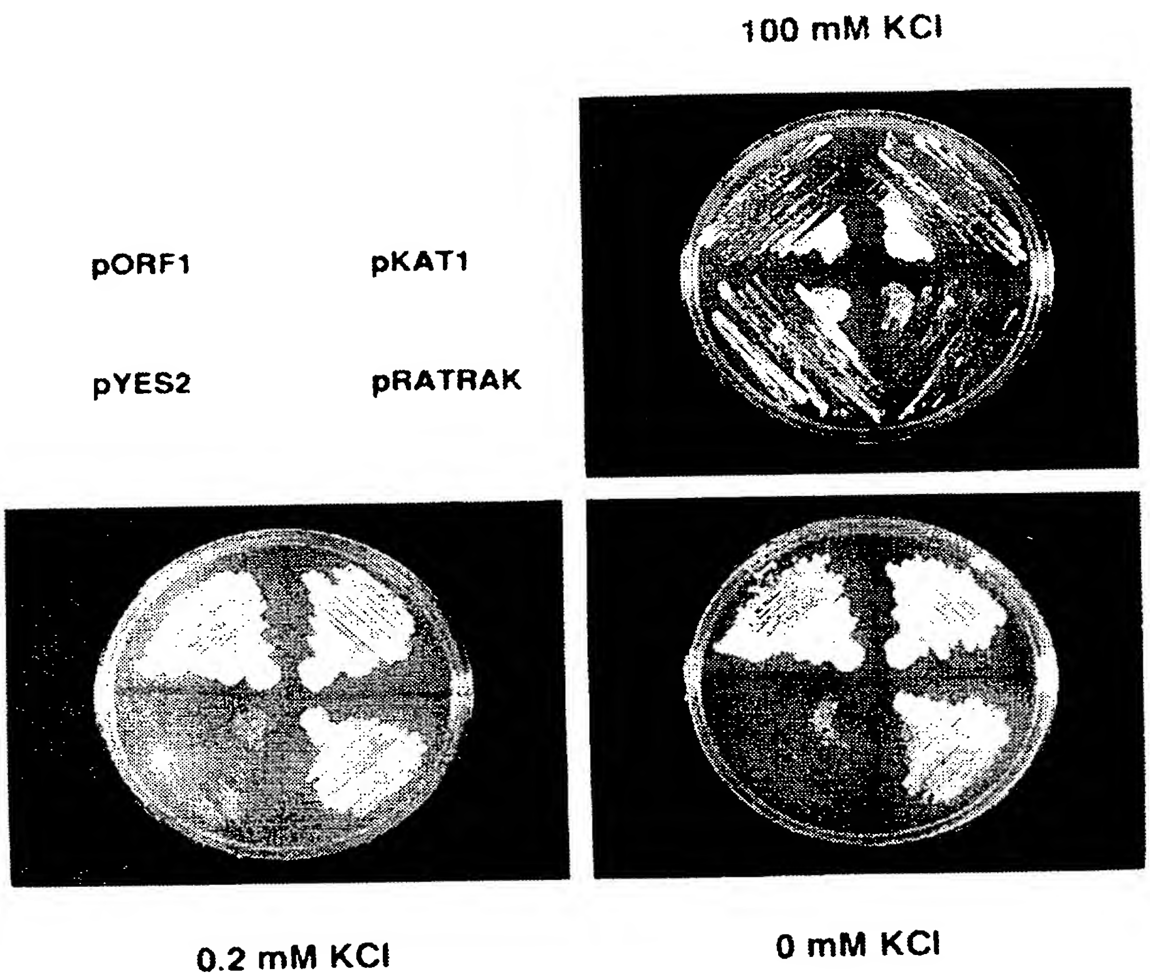


FIG. 7

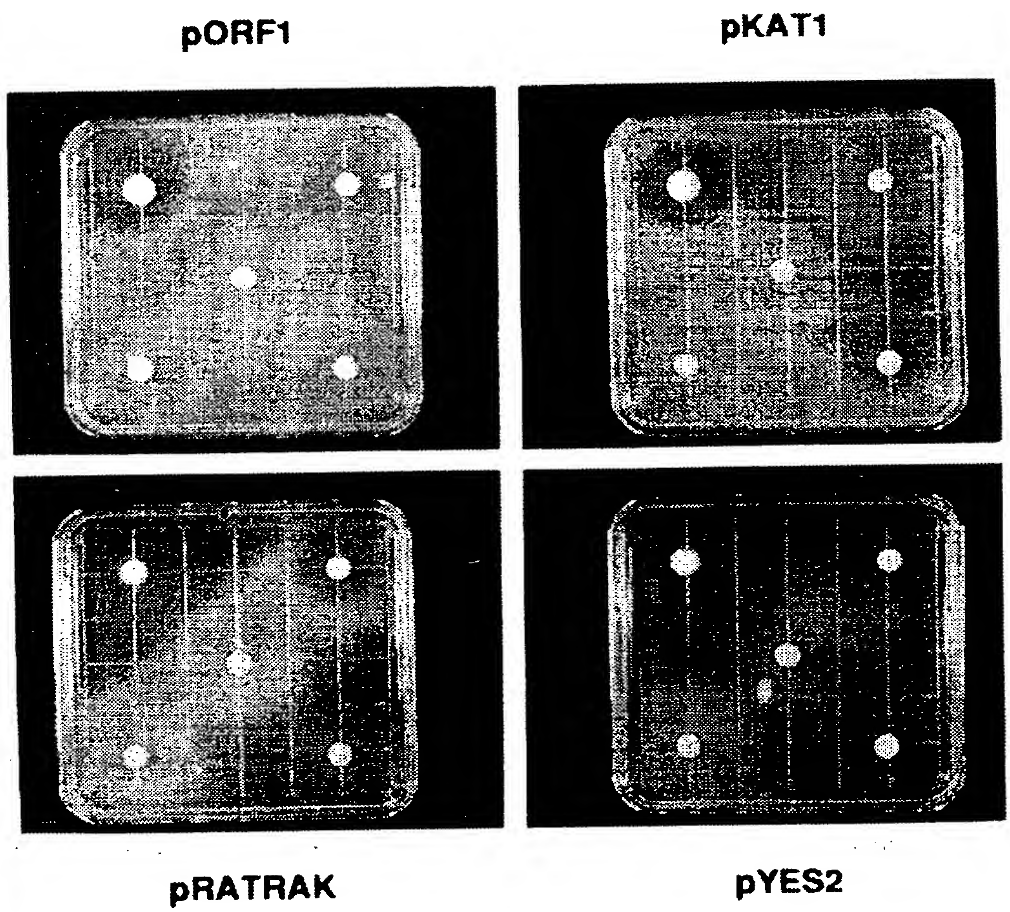


FIG. 8

12/13

```
                          G          10                                            20
Met Val Ile Ile Asn Arg Ser Asn Thr Tyr Ala Val Glu Gln Glu Ala Phe Pro Arg Asp Lys Tyr Asn Ile Val
ATG GTA ATA ATC AAC CGA TCG AAC ACC TAT GCC GTT GAG CAG GAA GCA TTT CCA AGA GAC AAG TAC AAT ATT GTC   75

                30                                      40                                      50
Tyr Trp Leu Val Ile Leu Val Gly Phe Gly Val Leu Leu Pro Trp Asn Met Phe Ile Thr Ile Ala Pro Glu Tyr
TAC TGG CTC GTC ATT CTT GTT GGA TTC GGA GTT CTT CTG CCA TGG AAT ATG TTC ATT ACT ATC GCC CCT GAG TAT  150

                                     60                                      70
Tyr Val Asn Tyr Trp Phe Lys Pro Asp Gly Val Glu Thr Trp Tyr Ser Lys Glu Phe Met Gly Ser Leu Thr Ile
TAT GTG AAT TAT TGG TTC AAA CCG GAT GGC GTG GAG ACA TGG TAT TCG AAA GAA TTC ATG GGA TCT TTG ACG ATT  225

                80          G                           90                                     100
Gly Ser Gln Leu Pro Asn Ala Ser Ile Asn Val Phe Asn Leu Phe Leu Ile Ile Ala Gly Pro Leu Ile Tyr Arg
GGC TCA CAA CTT CCA AAC GCA AGC ATT AAT GTT TTC AAC CTG TTC CTC ATT ATT GCT GGT CCC CTG ATC TAC CGC  300

                                    110                                     120
Val Phe Ala Pro Val Cys Phe Asn Ile Val Asn Leu Thr Ile Ile Leu Ile Leu Val Ile Val Leu Glu Pro Thr
GTC TTT GCT CCG GTT TGC TTC AAC ATC GTC AAC CTG ACA ATC ATT CTC ATC CTC GTC ATT GTT CTG GAG CCC ACT  375

                130                                     140                     G              150
Glu Asp Ser Met Ser Trp Phe Phe Trp Val Thr Leu Gly Met Ala Thr Ser Ile Asn Phe Ser Asn Gly Leu Tyr
GAA GAT TCC ATG TCC TGG TTT TTC TGG GTA ACT CTT GGA ATG GCG ACT TCA ATC AAT TTT AGC AAT GGG CTA TAT  450

                                    160                                     170
Glu Asn Ser Val Tyr Gly Val Gly Gly Asp Phe Pro His Thr Tyr Ile Gly Ala Leu Leu Ile Gly Asn Asn Ile
GAA AAC TCG GTT TAT GGA GTT GGT GGC GAT TTT CCG CAC ACC TAC ATT GGC GCT CTC TTG ATT GGA AAC AAC ATT  525

                180                                     190                                    200
Cys Gly Leu Leu Ile Thr Val Val Lys Ile Gly Val Thr Tyr Phe Leu Asn Asp Glu Pro Lys Leu Val Ala Ile
TGC GGA TTG CTG ATA ACG GTT GTG AAA ATC GGA GTG ACC TAT TTT CTG AAT GAT GAG CCT AAA CTT GTT GCA ATC  600

                                    210                                     220
Val Tyr Phe Gly Ile Ser Leu Val Ile Leu Leu Val Cys Ala Ile Ala Leu Phe Phe Ile Thr Lys Gln Asp Phe
GTC TAT TTC GGC ATA TCG TTG GTG ATC CTT CTG GTG TGT GCA ATT GCA CTT TTC TTT ATC ACA AAG CAA GAT TTC  675
```

FIG. 9A

```
                    230                                            240                                                  250
Tyr His Tyr His His Gln Lys Gly Met Glu Ile Arg Glu Lys Ala Glu Thr Asp Arg Pro Ser Pro Ser Ile Leu
TAC CAC TAT CAC CAT CAA AAA GGA ATG GAA ATT CGC GAA AAG GCG GAA ACC GAC AGA CCG TCT CCA TCC ATT CTT    750

                                    260                                     270
Trp Thr Thr Phe Thr Asn Cys Tyr Gly Gln Leu Phe Asn Val Trp Phe Cys Phe Ala Val Thr Leu Thr Ile Phe
TGG ACC ACA TTC ACA AAC TGT TAT GGG CAA CTC TTC AAT GTT TGG TTC TGC TTT GCC GTT ACT CTC ACA ATC TTC    825

                280                                     290                                     300
Pro Val Met Met Thr Val Thr Thr Arg Gly Asp Ser Gly Phe Leu Asn Lys Ile Met Ser Glu Asn Asp Glu Ile
CCT GTT ATG ATG ACC GTT ACC ACT CGT GGA GAT TCC GGC TTC CTA AAC AAA ATT ATG TCT GAA AAC GAT GAA ATC    900

                                    310                                     320
Tyr Thr Leu Leu Thr Ser Phe Leu Val Phe Asn Leu Phe Ala Ala Ile Gly Ser Ile Val Ala Ser Lys Ile His
TAC ACT TTG CTC ACA AGT TTC CTC GTC TTC AAT TTG TTC GCT GCG ATT GGA TCC ATA GTT GCT TCC AAG ATT CAC    975

                330                                     340                                     350
Trp Pro Thr Pro Arg Tyr Leu Lys Phe Ala Ile Ile Leu Arg Ala Leu Phe Ile Pro Phe Phe Phe Phe Cys Asn
TGG CCG ACA CCC CGT TAC CTC AAA TTT GCC ATA ATC TTG CGT GCT CTT TTC ATT CCA TTC TTC TTC TTC TGC AAC   1050

                                    360                                     370
Tyr Arg Val Gln Thr Arg Ala Tyr Pro Val Phe Phe Glu Ser Thr Asp Ile Phe Val Ile Gly Gly Ile Ala Met
TAT CGT GTC CAG ACG CGT GCT TAT CCT GTT TTC TTT GAG TCT ACT GAC ATT TTT GTG ATT GGT GGA ATT GCC ATG   1125

                380                                     390                                     400
Ser Phe Ser His Gly Tyr Leu Ser Ala Leu Ala Met Gly Tyr Thr Pro Asn Val Val Pro Ser His Tyr Ser Arg
TCT TTT TCA CAT GGA TAC CTC AGC GCT CTG GCA ATG GGA TAC ACT CCA AAC GTC GTG CCA TCT CAC TAC TCA AGA   1200

                                    410                                     420
Phe Ala Ala Gln Leu Ser Val Cys Thr Leu Met Val Gly Leu Leu Thr Gly Gly Leu Trp Pro Val Val Ile Glu
TTT GCC GCT CAG CTT TCC GTT TGC ACT CTT ATG GTT GGC CTT CTC ACC GGT GGC CTG TGG CCC GTT GTT ATT GAG   1275

                430                 434
His Phe Val Asp Lys Pro Ser Ile Leu
CAC TTC GTG GAC AAG CCA AGT ATC TTA TAA ATATTTATAGCATTAGAGTATACTTGTTATATGTTGTTTTTATTAAGCTGTGGAATAAA   1364

ATAATTATTAAAAAAAAAAAAAAA   1388
```

13/13

FIG. 9B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US 95/14364

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K14/705 C12N15/81 C12N1/19 C12Q1/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIOPHYS J, 63 (5). 1992. 1406-1411., MCCORMACK K ET AL 'TANDEM LINKAGE OF SHAKER POTASSIUM CHANNEL SUBUNITS DOES NOT ENSURE THE STOICHIOMETRY OF EXPRESSED CHANNELS' see the whole document ---	1
X	JOURNAL OF NEUROSCIENCE, 13 (11). 1993. 4669-4679., ZHONG Y ET AL 'Modulation of different K+ currents in Drosophila: A hypothetical role for the eag subunit in multimeric K+ channels' see the whole document ---	1

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 March 1996	27 MARCH 1996

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gurdjian, D

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	NATURE, vol. 345, 1990 pages 530-4, E.Y.ISACOFF ET AL. 'Evidence for the formation of heteromultimeric potassium channels in Xenopus oocytes ' see the whole document	1 20-26
X Y	NATURE, vol. 368, March 1994 pages 32-38, R. WILSON ET AL. '2.2 mb of contigous nucleotide sequence form chromosome III of c.elegans' see abstract; table 2	1 20-26
Y	SCIENCE , vol. 256, 1992 pages 663-5, H.SENTENAC ET AL. 'Cloning and expression in yeast of a plant potassium ion transport system' cited in the application see the whole document	20-26
Y	EP,A,0 615 976 (AMERICAN CYANAMID CO) 21 September 1994 see the whole document	20-26
A	PROC NATL ACAD SCI U S A, 86 (12). 1989. 4372-4376., KAMB A ET AL 'IDENTIFICATION OF GENES FROM PATTERN FORMATION TYROSINE KINASE AND POTASSIUM CHANNEL FAMILIES BY DNA AMPLIFICATION' see the whole document	10,12
A	US,A,5 356 775 (HEBERT STEVEN C ET AL) 18 October 1994 see the whole document	1,10-13
A	NATURE, vol. 362, 1993 pages 127-133, Y.KUBO ET AL. 'Primary structure and functional expression of a mouse inward rectifier potassium channel' cited in the application see the whole document	10-13

-/--

4

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	NATURE (LONDON), 376 (6542). 1995. 690-695., KETCHUM K A ET AL 'A new family of outwardly rectifying potassium channel proteins with two pore domains in tandem' see the whole document -----	1-13

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 23-26
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 23-26 refer, at least partially as far it concerns a medical method, to a method of treatment of the human or animal body, the search has been carried out and has been based on the alleged effects of the composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Info ...tion on patent family members

International ...plication No
PCT/US 95/14364

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0615976	21-09-94	CA-A- 2112445 JP-A- 6253849	01-07-94 13-09-94
US-A-5356775	18-10-94	NONE	

SC galactose, 100 mM KCl

SC glucose, 0mM KCl

SC galactose, 0 mM KCl

SC glucose, 100 mM KCl

FIG. 1

```
5'...ACGCGATCGCCGCGAGTGTATATTTTTTTTTTAGCTCAGTCTTCAGTGTTTCGCGATTCTCTTTAAAAGAAAAAAAAAATAATAAGTCAA
AACTACAAACCACACAGCGAAAGGCGAAAGCAACGGTTCCTGCGAGTGTTTATTTTTTTTTTCAACAATTTTTGATCGTAGTGCGACAATCCGTCGAGC      -1
                                      10                        M1       20
Met Ser Pro Asn Arg Trp Ile Leu Leu Leu Ile Phe Tyr Ile Ser Tyr Leu Met Phe Gly Ala Ala Ile Tyr Tyr
ATG TCG CCG AAT CGA TGG ATC CTG CTG CTC ATC TTC TAC ATA TCC TAC CTG ATG TTC GGG GCG GCA ATC TAT TAC    75
                                                 40                                              50
                    30
His Ile Glu His Gly Glu Glu Lys Ile Ser Arg Ala Glu Gln Arg Lys Ala Gln Ile Ala Ile Asn Glu Tyr Leu
CAT ATT GAG CAC GGC GAG GAG AAG ATA TCG CGC GCC GAA CAG CGC AAG GCG CAA ATT GCA ATC AAC GAA TAT CTG    150
                            G   60                                     70
Leu Glu Glu Leu Gly Asp Lys Asn Thr Thr Thr Gln Asp Glu Ile Leu Gln Arg Ile Ser Asp Tyr Cys Asp Lys
CTG GAG GAG CTG GGC GAC AAG AAT ACG ACC ACA CAG GAT GAG ATT CTT CAA CGG ATC TCG GAT TAC TGT GAC AAA    225
                    80                                 90                                        100
Pro Val Thr Leu Pro Pro Thr Tyr Asp Asp Thr Pro Tyr Thr Trp Thr Phe Tyr His Ala Phe Phe Phe Ala Phe
CCG GTT ACA TTG CCG CCG ACA TAT GAT GAT ACG CCC TAC ACG TGG ACC TTC TAC CAT GCC TTC TTC TTC GCC TTC    300
                                    110                                    120
        B5-1
Thr Val Cys Ser Thr Val Gly Tyr Gly Asn Ile Ser Pro Thr Thr Phe Ala Gly Arg Met Ile Met Ile Ala Tyr
ACC GTT TGC TCC ACG GTG GGA TAT GGG AAT ATA TCG CCA ACC ACC TTC GCC GGA CGG ATG ATC ATG ATC GCG TAT    375
                130     M2                                 140                                   150
Ser Val Ile Gly Ile Pro Val Asn Gly Ile Leu Phe Ala Gly Leu Gly Glu Tyr Phe Gly Arg Thr Phe Glu Ala
TCG GTG ATT GGC ATC CCC GTC AAT GGT ATC CTC TTT GCC GGC CTC GGC GAA TAC TTT GGA CGT ACG TTT GAA GCG    450
                                    160                                    170
Ile Tyr Arg Arg Tyr Lys Lys Tyr Lys Met Ser Thr Asp Met His Tyr Val Pro Pro Gln Leu Gly Leu Ile Thr
ATC TAC AGA CGC TAC AAA AAG TAC AAG ATG TCC ACG GAT ATG CAC TAT GTC CCG CCG CAG CTG GGA TTG ATC ACC    525
                180     M3                                 190                                   200
Thr Val Val Ile Ala Leu Ile Pro Gly Ile Ala Leu Phe Leu Val Leu Pro Cys Val Gly Val His Leu Leu Arg
ACG GTG GTG ATT GCC CTG ATT CCG GGA ATA GCT CTC TTC CTG GTG CTG CCC TGC GTG GGT GTT CAC CTA CTT CGA    600
                                    210                        B5-2                  220
Glu Leu Gly Leu Ser Ser Ile Ser Leu Tyr Tyr Ser Tyr Val Thr Thr Thr Thr Ile Gly Phe Gly Asp Tyr Val
GAA CTG GGC CTA TCT TCC ATC TCG CTG TAC TAC AGC TAT GTG ACC ACC ACA ACA ATT GGA TTC GGT GAC TAT GTG    675
                230                                        240                                   250
Pro Thr Phe Gly Ala Asn Gln Pro Lys Glu Phe Gly Gly Trp Phe Val Val Tyr Gln Ile Phe Val Ile Val Trp
CCC ACA TTT GGA GCC AAC CAG CCC AAG GAG TTC GGC GGC TGG TTC GTG GTC TAT CAG ATC TTT GTG ATC GTG TGG    750
                                    260                                    270
    M4
Phe Ile Phe Ser Leu Gly Tyr Leu Val Met Ile Met Thr Phe Ile Thr Arg Gly Leu Gln Ser Lys Lys Leu Ala
TTC ATC TTC TCG CTG GGA TAT CTT GTG ATG ATC ATG ACA TTT ATC ACT CGG GGC CTC CAG AGC AAG AAG CTG GCA    825
                280                                        290                                   300
Tyr Leu Glu Gln Gln Leu Ser Ser Asn Leu Lys Ala Thr Gln Asn Arg Ile Trp Ser Gly Val Thr Lys Asp Val
TAC CTG GAG CAG CAG TTG TCC TCC AAC CTG AAG GCC ACA CAG AAT CGC ATC TGG TCT GGC GTC ACC AAG GAT GTG    900
                                    310                                    320
Gly Tyr Leu Arg Arg Met Leu Asn Glu Leu Tyr Ile Leu Lys Val Lys Pro Val Tyr Thr Asp Val Asp Ile Ala
GGC TAC CTC CGG CGA ATG CTC AAC GAG CTG TAC ATC CTC AAA GTG AAG CCT GTG TAC ACC GAT GTA GAT ATC GCC    975
```

FIG. 2A

```
                330                                     340                     350
Tyr Thr Leu Pro Arg Ser Asn Ser Cys Pro Asp Leu Ser Met Tyr Arg Val Glu Pro Ala Pro Ile Pro Ser Arg
TAC ACA CTG CCA CGT TCC AAT TCG TGT CCG GAT CTG AGC ATG TAC CGC GTG GAG CCG GCT CCC ATT CCC AGC CGG   1050
                                    360                                     370
Lys Arg Ala Phe Ser Val Cys Ala Asp Met Val Gly Ala Gln Arg Glu Ala Gly Met Val His Ala Asn Ser Asp
AAG AGG GCA TTC TCC GTG TGC GCC GAC ATG GTT GGC GCC CAA AGG GAG GCG GGC ATG GTA CAC GCC AAT TCC GAT   1125
                380                                     390                                     400
Thr Asp Leu Thr Lys Leu Asp Arg Glu Lys Thr Phe Glu Thr Ala Glu Ala Tyr His Gln Thr Thr Asp Leu Leu
ACG GAT CTA ACC AAA CTG GAT CGC GAG AAG ACA TTC GAG ACG GCG GAG GCG TAC CAC CAG ACC ACC GAT TTG CTG   1200
                                    410                                     420
Ala Lys Val Val Asn Ala Leu Ala Thr Val Lys Pro Pro Pro Ala Glu Gln Glu Asp Ala Ala Leu Tyr Gly Gly
GCC AAG GTG GTC AAC GCA CTG GCC ACG GTG AAG CCA CCG CCG GCG GAA CAG GAA GAT GCG GCT CTC TAT GGT GGC   1275
                430                                     440                                     450
Tyr His Gly Phe Ser Asp Ser Gln Ile Leu Ala Ser Glu Trp Ser Phe Ser Thr Val Asn Glu Phe Thr Ser Pro
TAT CAT GGC TTC TCC GAC TCC CAG ATC CTG GCC AGC GAA TGG TCG TTC TCG ACG GTC AAC GAG TTC ACA TCA CCG   1350
                                    460                                     470
Arg Arg Pro Arg Ala Arg Ala Cys Ser Asp Phe Asn Leu Glu Ala Pro Arg Trp Gln Ser Glu Arg Pro Leu Arg
CGA CGT CCA AGA GCA CGT GCC TGC TCC GAT TTC AAT CTG GAG GCA CCT CGC TGG CAG AGC GAG AGG CCA CTG CGT   1425
                480                                     490                                     500
Ser Ser His Asn Glu Trp Thr Trp Ser Gly Asp Asn Gln Gln Ile Gln Glu Ala Phe Asn Gln Arg Tyr Lys Gly
TCG AGC CAC AAC GAA TGG ACA TGG AGC GGC GAC AAC CAG CAG ATC CAG GAG GCA TTC AAC CAG CGC TAC AAG GGA   1500
                                    510                                     520
Gln Gln Arg Ala Asn Gly Ala Ala Asn Ser Thr Met Val His Leu Glu Pro Asp Ala Leu Glu Glu Gln Leu Arg
CAG CAG CGT GCC AAC GGA GCA GCC AAC TCG ACC ATG GTC CAT CTG GAG CCG GAT GCT TTG GAG GAG CAG CTG AGA   1575
                530                                     540                                     550
Asn Asn His Arg Val Pro Val Ala Ser Arg Ser Ser Pro Cys Arg Met Val Cys Asp Val Cys Phe Pro Ser Arg
AAC AAT CAC CGG GTG CCG GTC GCG TCA AGA AGT TCT CCA TGC CGG ATG GTC TGC GAC GTC TGT TTC CCT TCC AGA   1650
                                    560                                     570
Arg Ser Thr Pro Arg Arg Ile Trp Ser Ala Ser Cys Pro Trp Ser Arg Tyr Pro Arg Val Ser Ser Arg Arg Lys
AGA AGC ACC CCT CGC AGG ATC TGG AGC GCA AGT TGT CCG TGG TCT CGG TAC CCG AGG GTG TCA TCT CGC AGG AAG   1725
                580                                     590                                     600
Pro Asp Pro Arg Trp Thr Thr Thr Ser Thr Arg Ser Arg Arg Pro Pro Val Asn Pro Ile Cys Ala Thr Asp Ala
CCA GAT CCC CGC TGG ACT ACT ACA TCA ACA CGG TCA CGG CGG CCT CCA GTC AAT CCT ATT TGC GCA ACG GAC GCG   1800
                                    610                         618
Val Arg His Arg Pro Ser Asn Arg Met Ala Ala Trp Pro Ala Ala Ala Ala Gly
GTC CGC CAC CGC CCT TCG AAT CGA ATG GCA GCT TGG CCA GCG GCG GCG GCG GGC TAA CGAACATGGGCTTCCAGATGGAG   1880

GATGGAGCAACCCCGCCATCGGCATTGGGCGGTGGAGCCTATCAACGCAAGGCGGCTGCTGGCAAGCGCCGACGCGAGAGCATCTACACCCAGAATCAA
GCCCCATCCGCTCGCCGGGGCAGCATGTATCCGCCGACCGCGCACGCCTTGGCCCAGATGCAGATGCGACGCGGCAGCTTGGCAACCAGTGGCTCTGGA
TCGGCGGCCATGGCGGCAGTGGCCGCGCGTCGTGGCAGCCTCTTCCCAGCTACAGCATCGGCATCATCGCTGACCTCTGCTCCGCGCCGAAGCAGCATA
TTCTCGGTTACCTCCGAAAAGGATATGAATGTGCTGGAGCAGACGACCATTGCGGATCTGATTCGTGCGCTCGAG . . .3'
```

FIG. 2B

```
                                            10                                          20
Met Ser Asp Gln Leu Phe Val Ala Phe Glu Lys Tyr Phe Leu Thr Ser Asn Glu Val Lys
ATG TCC GAT CAG CTG TTT GTC GCA TTT GAG AAG TAT TTC TTG ACG AGT AAC GAG GTC AAG    60
                                    30                                      40
Lys Asn Ala Ala thr Glu Thr Trp Thr Phe Ser Ser Ser Ile Phe Phe Ala Val Thr Val
AAG AAT GCA GCA ACG GAG ACA TGG ACA TTT TCA TCG TCC ATT TTC TTT GCC GTA ACC GTC   120
    H5-1                                50                                      60
Val Thr Thr Ile Gly Tyr Gly Asn Pro Val Pro Val Thr Asn Ile Gly Arg Ile Trp Cys
GTC ACT ACC ATC GGA TAC GGT AAT CCA GTT CCA GTG ACA AAC ATT GGA CGG ATA TGG TGT   180
                        M2              70                                      80
Ile Leu Phe Ser Leu Leu Gly Ile Pro Leu Thr Leu Val Thr Ile Ala Asp Leu Ala Gly
ATA TTG TTC TCC TTG CTT GGA ATA CCT CTA ACA CTG GTT ACC ATC GCT GAC TTG GCA GGT   240
                                    90                                      100
Lys Phe Leu Ser Glu His Leu Val Trp Leu Tyr Gly Asn Tyr Leu Lys Leu Lys Tyr Leu
AAA TTC CTA TCT GAA CAT CTT GTT TGG TTG TAT GGA AAC TAT TTG AAA TTA AAA TAT CTC   300
                                    110                                     120
Ile Leu Ser Arg His Arg Lys Glu Arg Arg Glu His Val Cys Glu His Cys His Ser His
ATA TTG TCA CGA CAT CGA AAA GAA CGG AGA GAG CAC GTT TGT GAG CAC TGT CAC AGT CAT   360
                                    130                                     140
Gly Met Gly His Asp Met Asn Ile Glu Glu Lys Arg Ile Pro Ala Phe Leu Val Leu Ala
GGA ATG GGG CAT GAT ATG AAT ATC GAG GAG AAA AGA ATT CCT GCA TTC CTG GTA TTA GCT   420
            M3                          150                                     160
Ile Leu Ile Val Tyr Thr Ala Phe Gly Gly Val Leu Met Ser Lys Leu Glu Pro Trp Ser
ATT CTG ATA GTA TAT ACA GCG TTT GGC GGT GTC CTA ATG TCA AAA TTA GAG CCG TGG TCT   480
```

FIG. 3A

```
                                          170        K5-2                             180
Phe Phe Thr Ser Phe Tyr Trp Ser Phe Ile Thr Met Thr Thr Val Gly Phe Gly Asp Leu
TTC TTC ACT TCA TTC TAC TGG TCC TTC ATT ACA ATG ACT ACT GTC GGG TTT GGC GAC TTG     540
                                    190                                     200
Met Pro Arg Arg Asp Gly Tyr Met Tyr Ile Ile Leu Leu Tyr Ile Ile Leu Gly Lys Phe
ATG CCC AGA AGG GAC GGA TAC ATG TAT ATC ATA TTG CTC TAT ATC ATT TTA GGT AAA TTT     600
                                    210                             M4      220
Ser Met Lys Lys Lys Gln Lys Phe Lys Ile Phe Leu Gly Leu Ala Ile Thr Thr Met Cys
TCA ATG AAA AAA AAA CAA AAA TTC AAA ATA TTT TTA GGT CTT GCA ATA ACT ACA ATG TGC     660
                                    230                                     240
Ile Asp Leu Val Gly Val Gln Tyr Ile Arg Lys Ile His Tyr Phe Gly Arg Lys Ile Gln
ATT GAT TTG GTA GGA GTA CAG TAT ATT CGA AAG ATT CAT TAT TTC GGA AGA AAA ATT CAA     720
                                    250                                     260
Asp Ala Arg Ser Ala Leu Ala Val Val Gly Gly Lys Val Val Leu Val Ser Glu Leu Tyr
GAC GCT AGA TCT GCA TTG GCG GTT GTA GGA GGA AAG GTA GTC CTT GTA TCA GAA CTC TAC     780
                                    270                                     280
Ala Asn Leu Met Gln Lys Arg Ala Arg Asn Met Ser Arg Glu Ala Phe Ile Val Glu Asn
GCA AAT TTA ATG CAA AAG CGA GCT CGT AAC ATG TCC CGA GAA GCT TTT ATA GTG GAG AAT     840
                                    290                                     300
Leu Tyr Val Ser Lys His Ile Ile Pro Phe Ile Pro Thr Asp Ile Arg Cys Ile Arg Tyr
CTC TAT GTT TCC AAA CAC ATC ATA CCA TTC ATA CCA ACT GAT ATC CGA TGT ATT CGA TAT     900
                                    310                                     320
Ile Asp Gln Thr Ala Asp Ala Ala Thr Ile Ser Thr Ser Ser Ser Ala Ile Asp Met Gln
ATT GAT CAA ACT GCC GAT GCT GCT ACC ATT TCC ACG TCA TCG TCT GCA ATT GAT ATG CAA     960
                                    330                     336
Ser Cys Arg Phe Cys His Ser Arg Tyr Ser Leu Asn Arg Ala Phe Lys
AGT TGT AGA TTT TGT CAT TCA AGA TAT TCT CTC AAT CGT GCA TTC AAA TAG                1011
```

FIG. 3B

```
Ce orf1     ---------- ---------- ---------- ---------- ----------
Dm orf1     MSPNRWILLL IFYISYLMFG AAIYYHIEHG EEKISRAEQR KAQIAINEYL   50
Consensus   .......... .......... .......... .......... ..........   50

Ce orf1     ---------- --MSDQLFVA FEKYFLTSNE VKKNAATETW TFSSSIFFAV   38
Dm orf1     LEELGDKNTT TQDEILQRIS DYCDKPVTLP PTYDDTPYTW TFYHAFFFAF  100
Consensus   .......... .......... .......... ........TW TF....FFA.  100

Ce orf1     TVVTTIGYGN PVPMTNIGRI WCILFSLLGI PLTLVTIADL AGKFLSEHLV   88
Dm orf1     TVCSTMGYGN ISPTTFAGRM IMIAYSVIGI PVNGILFAAL ----------  140
Consensus   TV..T.GYGN ..P.T..GR. ..I..S..GI P......A.L ..........  150

Ce orf1     WLYGNYLKLK YLILSRHRKE RREHVCEHCH SHQMGHQMNI EEKRIPAFLV  138
Dm orf1     ---GEYPGRT FEAIYRRYKK YKMSTDMHYV PPQLGLITTV VIALIPGIAL  187
Consensus   ...G.Y.... .....R..K. .......H.. ....G..... ....IP....  200

Ce orf1     LAILIVYTAF GGVLMSKLEP WSFFTSFYWS FIIMTTVGFG DLMPRRDGYM  188
Dm orf1     FLVLPCVGVH LLRELGLSS- ----ISLYMS YVIITTIGFG DYVPT-FGAN  231
Consensus   ...L...... .......... .....S.Y.S ..I.TT.GFG D..P...G..  250

Ce orf1     YIILLYIILG KFSMKKKQHF KIFLGLAITT MCIDLVGMQY IRKIHYFGRK  238
Dm orf1     QPKEFGGWFV VYQIFVIVWF IFSLGYLVMI MTFITFGLQS KKLAYLEQQL  281
Consensus   .......... .........F ...LG..... M.....G.Q. ..........  300

Ce orf1     IQDARSALAV VGGKVVLVSE LYANLMQKRA RNMSREAFIV ENLYVSKHII  288
Dm orf1     SSNLKATQNR IWSGVTKDVG YLRRMLNELY ILKVKPVYTD VDIAYTLPRS  331
Consensus   .......... ....V..... .......... .......... ..........  350

Ce orf1     PFIPTDIRCI -RYIDQTADA ATISTSSSAI DMQSCRFCHS RYSLNRAFKK  337
Dm orf1     NSCPDLSMYR VEPAPIPSRK RAFSVCADMV GAQREAGMVH ANSDTDLTKL  381
Consensus   ...P...... .......... ...S...... ..Q....... ..S.....K.  400

Ce orf1     ---------- ---------- ---------- ---------- ----------  337
Dm orf1     DREKTFETAE AYHQTTDLLA KVVNALATVK PPPAECEDAA LYGGYHGFSD  431
Consensus   .......... .......... .......... .......... ..........  450

C  orf1     ---------- ---------- ---------- ---------- ----------  337
Dm orf1     SQILASEWSF STVNEFTSPR RPRARACSDF MLEAPRWQSE RPLRSSHNEW  481
Consensus   .......... .......... .......... .......... ..........  500
```

FIG. 4

```
mIRK        AFLFSIETQTTIGYGFRCVTDECP      {G,A,S,T},  {D,E}
hROMK1      AFLFSLETQVTIGYGFRCVTEQCA      {N,Q},  {K,R,H}
rGIRK1      AFLFFIETEATIGYGYRYITDHCP      {F,Y,W}={I,L,M,V}
            ||.|..   .|.|||    |   .
Dm H5-1     AFFFAFTVCSTVGYGNISPTTFAG
            ||.||.|  ||||||||.||||||
Shak        AFWWAVVTMTTVGYGDMTPVGFWG
Shal        AFWYTIVTMTTLGYGDMVPETIAG
Shab        AFWWAGITMTTVGYGDICPTTALG
Shaw        GLWWALVTMTTVGYGDMAPKTYIG
Eag         ALYFTMTCMTSVGFGNVAAETDNE
Slo         CVYFLIVTMSTVGYGDVYCETVLG
            .|| |.|| |||||||.|||  .|
Dm H5-2     SLYTSYVTTTTIGFGDYVPTFGAN

Dm H5-1     AFFFAFTVCSTVGYGNISPTTFAG
Ce 5-1      SIFFAVTVVTTIGYGNPVPVTNTG
Dm H5-2     SLYTSYVTTTTIGFGDYVPTFGAN
Ce H5-2     SFYWSFITMTTVGFGDLMPRRDGY
```

FIG. 5A

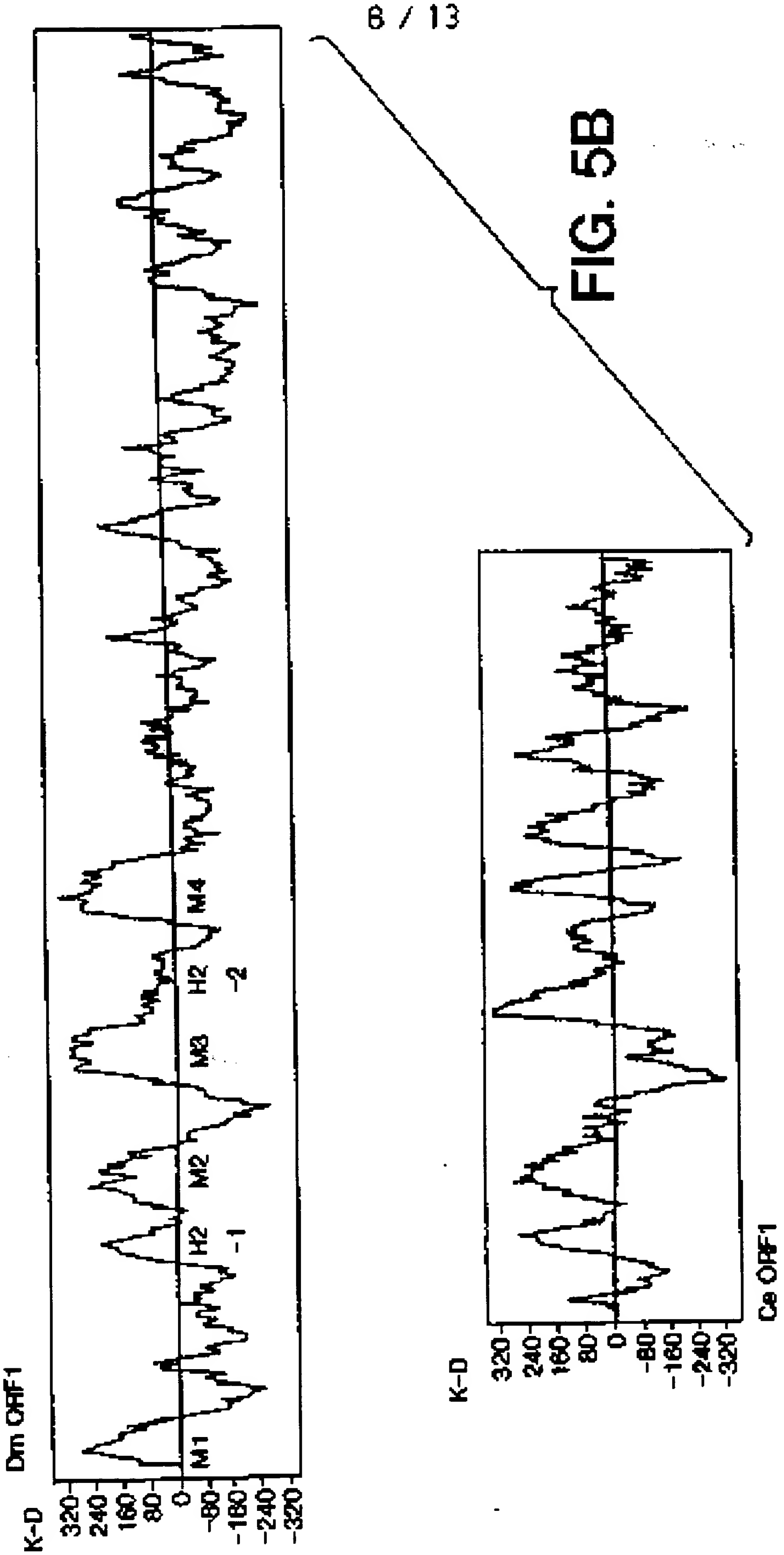

K-D
320
240
160
80
0
-80
-160
-240
-320
Dm ORF1
M1
H2
-1
M2
M3
H2
-2
M4
K-D
320
240
160
80
0
-80
-160
-240
-320
Ce ORF1

FIG. 5B

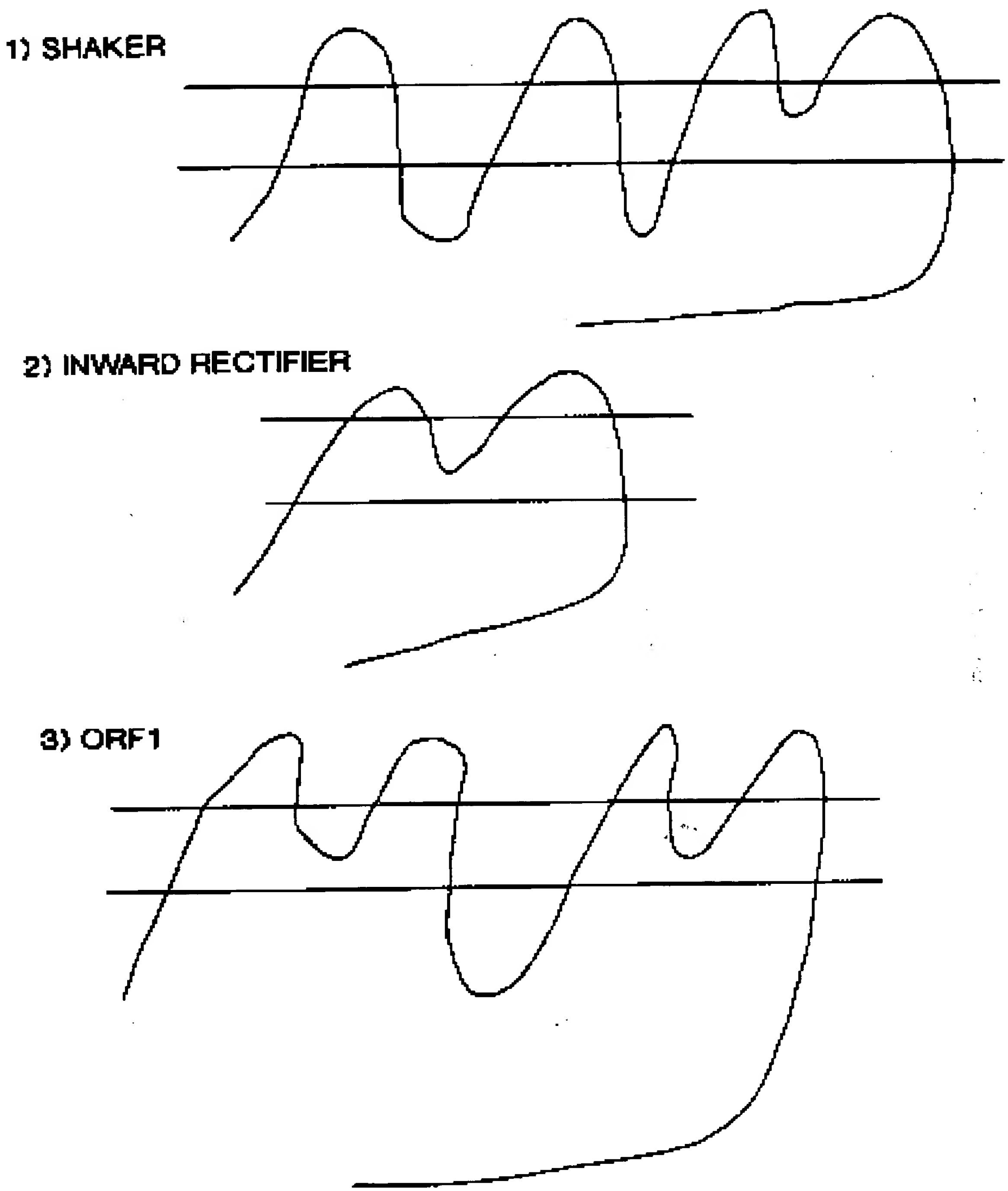


FIG. 6

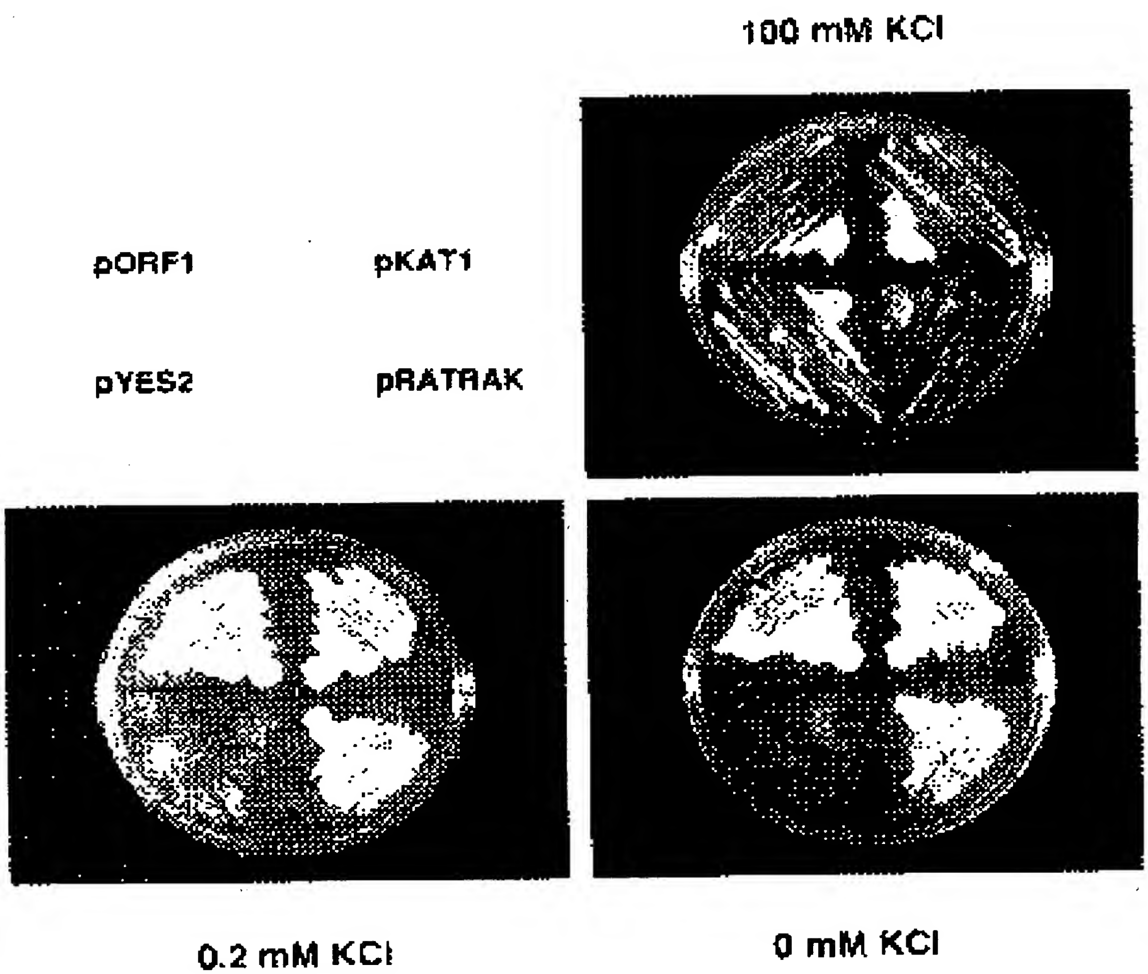


FIG. 7

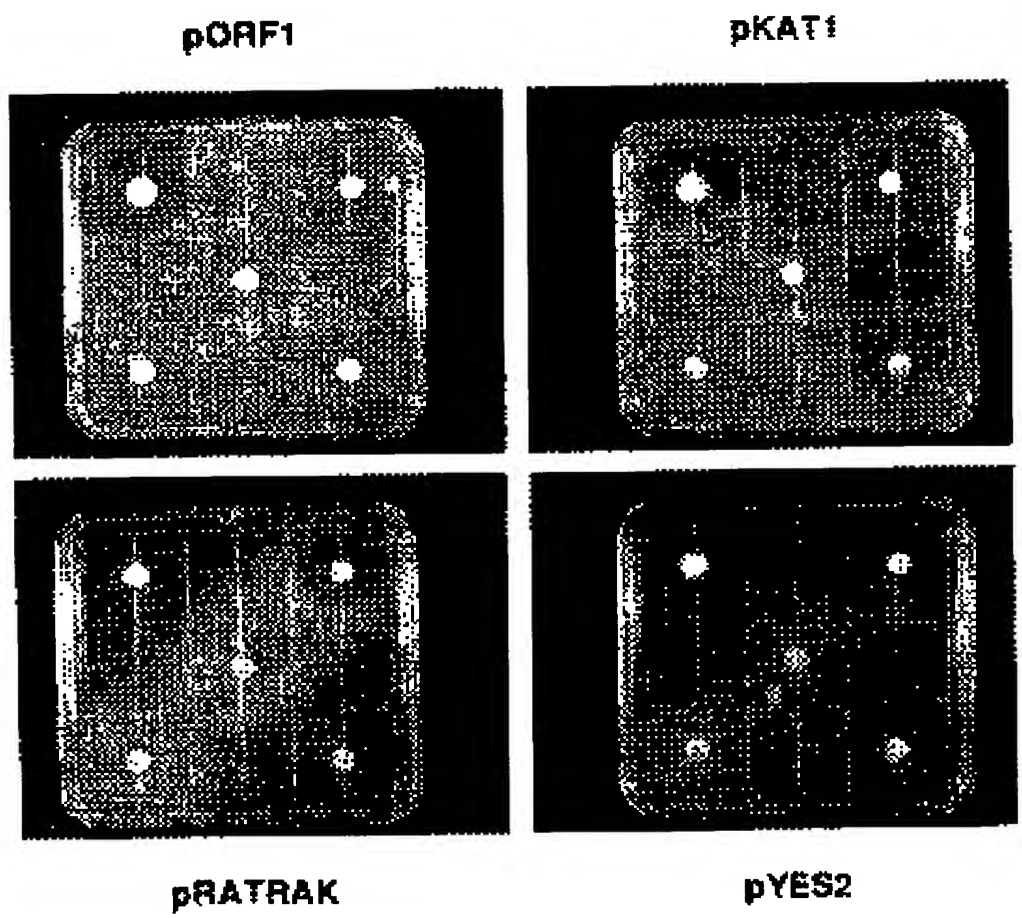


FIG. 8

12/13

```
                        G          10                                      20
Met Val Ile Ile Asn Arg Ser Asn Thr Tyr Ala Val Glu Gln Glu Ala Phe Pro Arg Asp Lys Tyr Asn Ile Val
ATG GTA ATA ATC AAC CGA TCG AAC ACC TAT GCC GTT GAG CAG GAA GCA TTT CCA AGA GAC AAG TAC AAT ATT GTC   75

                30                                      40                                      50
Tyr Trp Leu Val Ile Leu Val Gly Phe Gly Val Leu Leu Pro Trp Asn Met Phe Ile Thr Ile Ala Pro Glu Tyr
TAC TGG CTC GTC ATT CTT GTT GGA TTC GGA GTT CTT CTG CCA TGG AAT ATG TTC ATT ACT ATC GCC CCT GAG TAT   150

                                   60                                      70
Tyr Val Asn Tyr Trp Phe Lys Pro Asp Gly Val Glu Thr Trp Tyr Ser Lys Glu Phe Met Gly Ser Leu Thr Ile
TAT GTG AAT TAT TGG TTC AAA CCG GAT GGC GTG GAG ACA TGG TAT TCG AAA GAA TTC ATG GGA TCT TTG ACG ATT   225

                80          G                          90                                     100
Gly Ser Gln Leu Pro Asn Ala Ser Ile Asn Val Phe Asn Leu Phe Leu Ile Ile Ala Gly Pro Leu Ile Tyr Arg
GGC TCA CAA CTT CCA AAC GCA AGC ATT AAT GTT TTC AAC CTG TTC CTC ATT ATT GCT GGT CCC CTG ATC TAC CGC   300

                                   110                                     120
Val Phe Ala Pro Val Cys Phe Asn Ile Val Asn Leu Thr Ile Ile Leu Ile Leu Val Ile Val Leu Glu Pro Thr
GTC TTT GCT CCG GTT TGC TTC AAC ATC GTC AAC CTG ACA ATC ATT CTC ATC CTC GTC ATT GTT CTG GAG CCC ACT   375

                130                                     140                     G              150
Glu Asp Ser Met Ser Trp Phe Phe Trp Val Thr Leu Gly Met Ala Thr Ser Ile Asn Phe Ser Asn Gly Leu Tyr
GAA GAT TCC ATG TCC TGG TTT TTC TGG GTA ACT CTT GGA ATG GCG ACT TCA ATC AAT TTT AGC AAT GGG CTA TAT   450

                                   160                                     170
Glu Asn Ser Val Tyr Gly Val Gly Gly Asp Phe Pro His Thr Tyr Ile Gly Ala Leu Leu Ile Gly Asn Asn Ile
GAA AAC TCG GTT TAT GGA GTT GGT GGC GAT TTT CCG CAC ACC TAC ATT GGC GCT CTC TTG ATT GGA AAC AAC ATT   525

                180                                     190                                     200
Cys Gly Leu Leu Ile Thr Val Val Lys Ile Gly Val Thr Tyr Phe Leu Asn Asp Glu Pro Lys Leu Val Ala Ile
TGC GGA TTG CTG ATA ACG GTT GTG AAA ATC GGA GTG ACC TAT TTT CTG AAT GAT GAG CCT AAA CTT GTT GCA ATC   600

                                   210                                     220
Val Tyr Phe Gly Ile Ser Leu Val Ile Leu Leu Val Cys Ala Ile Ala Leu Phe Phe Ile Thr Lys Gln Asp Phe
GTC TAT TTC GGC ATA TCG TTG GTG ATC CTT CTG GTG TGT GCA ATT GCA CTT TTC TTT ATC ACA AAG CAA GAT TTC   675
```

FIG. 9A

13/13

```
                 230                                     240                                     250
Tyr His Tyr His His Gln Lys Gly Met Glu Ile Arg Glu Lys Ala Glu Thr Asp Arg Pro Ser Pro Ser Ile Leu
TAC CAC TAT CAC CAT CAA AAA GGA ATG GAA ATT CGC GAA AAG GCG GAA ACC GAC AGA CCG TCT CCA TCC ATT CTT   750

                                     260                                     270
Trp Thr Thr Phe Thr Asn Cys Tyr Gly Gln Leu Phe Asn Val Trp Phe Cys Phe Ala Val Thr Leu Thr Ile Phe
TGG ACC ACA TTC ACA AAC TGT TAT GGG CAA CTC TTC AAT GTT TGG TTC TGC TTT GCC GTT ACT CTC ACA ATC TTC   825

                 280                                     290                                     300
Pro Val Met Met Thr Val Thr Thr Arg Gly Asp Ser Gly Phe Leu Asn Lys Ile Met Ser Glu Asn Asp Glu Ile
CCT GTT ATG ATG ACC GTT ACC ACT CGT GGA GAT TCC GGC TTC CTA AAC AAA ATT ATG TCT GAA AAC GAT GAA ATC   900

                                     310                                     320
Tyr Thr Leu Leu Thr Ser Phe Leu Val Phe Asn Leu Phe Ala Ala Ile Gly Ser Ile Val Ala Ser Lys Ile His
TAC ACT TTG CTC ACA AGT TTC CTC GTC TTC AAT TTG TTC GCT GCG ATT GGA TCC ATA GTT GCT TCC AAG ATT CAC   975

                 330                                     340                                     350
Trp Pro Thr Pro Arg Tyr Leu Lys Phe Ala Ile Ile Leu Arg Ala Leu Phe Ile Pro Phe Phe Phe Phe Cys Asn
TGG CCG ACA CCC CGT TAC CTC AAA TTT GCC ATA ATC TTG CGT GCT CTT TTC ATT CCA TTC TTC TTC TTC TGC AAC   1050

                                     360                                     370
Tyr Arg Val Gln Thr Arg Ala Tyr Pro Val Phe Phe Glu Ser Thr Asp Ile Phe Val Ile Gly Gly Ile Ala Met
TAT CGT GTC CAG ACG CGT GCT TAT CCT GTT TTC TTT GAG TCT ACT GAC ATT TTT GTG ATT GGT GGA ATT GCC ATG   1125

                 380                                     390                                     400
Ser Phe Ser His Gly Tyr Leu Ser Ala Leu Ala Met Gly Tyr Thr Pro Asn Val Val Pro Ser His Tyr Ser Arg
TCT TTT TCA CAT GGA TAC CTC AGC GCT CTG GCA ATG GGA TAC ACT CCA AAC GTC GTG CCA TCT CAC TAC TCA AGA   1200

                                     410                                     420
Phe Ala Ala Gln Leu Ser Val Cys Thr Leu Met Val Gly Leu Leu Thr Gly Gly Leu Trp Pro Val Val Ile Glu
TTT GCC GCT CAG CTT TCC GTT TGC ACT CTT ATG GTT GGC CTT CTC ACC GGT GGC CTG TGG CCC GTT GTT ATT GAG   1275

                 430                     434
His Phe Val Asp Lys Pro Ser Ile Leu
CAC TTC GTG GAC AAG CCA AGT ATC TTA TAA ATATTTATAGCATTAGAGTATACTTGTTATATGTTGTTTTTATTAAGCTGTGGAATAAA   1364

ATAATTATTAAAAAAAAAAAAAAA   1388
```

FIG. 9B